

HUMAN DEHYDROGENASE GENE AND POLYPEPTIDE

DESCRIPTION OF THE DRAWINGS

SEQ NOS 1 and 2 show the nucleotide and amino acid sequences of human ANH401.
5 SEQ ID NO 3 shows the amino acid sequence of AF326966, a variant of ANH401, and SEQ ID NO 4 is the amino acid sequence of XM_048113, which lacks a substantial part of ANH401.

Fig. 1 is the alignment of the amino acid sequences of human ANH401 (SEQ ID NO 2), AF326966 (SEQ ID NO 3), and XM_048113 (SEQ ID NO 4).

DESCRIPTION OF THE INVENTION

The present invention relates to all facets of novel polynucleotides, the polypeptides they encode, antibodies and specific binding partners thereto, and their applications to research, diagnosis, drug discovery, therapy, clinical medicine, forensic science and
15 medicine, etc. The polynucleotides are expressed in angiogenesis and are therefore useful in variety of ways, including, but not limited to, as molecular markers for blood vessels and blood vessel formation, as drug targets, and for detecting, diagnosing, staging, monitoring, prognosticating, preventing, treating, and/or determining predisposition to diseases and conditions of the vascular system. The identification of specific genes, and groups of genes,
20 expressed in pathways physiologically relevant to angiogenesis permits the definition of functional and disease pathways, and the delineation of targets in these pathways which are useful in diagnostic, therapeutic, and clinical applications. The present invention also relates to methods of using the polynucleotides and related products (proteins, antibodies, etc.) in business and computer-related methods, e.g., advertising, displaying, offering, selling, etc.,
25 such products for sale, commercial use, licensing, etc.

Angiogenesis, the process of blood vessel formation, is a key event in many physiological processes that underlie normal and diseased tissue function. During ontogeny, angiogenesis is necessary to establish the network of blood vessels required for normal cell, tissue and organ development and maintenance. In the adult organism, the production of
30 new blood vessels is needed for organ homeostasis, e.g., in the cycling of the female endometrium, for blood vessel maturation during wound healing, and other processes

involved in the maintenance of organism integrity. It also is important in regenerative medicine, including, e.g., in promoting tissue repair, tissue engineering, and the growth of new tissues, inside and outside the body.

Not all angiogenesis is beneficial. Inappropriate and ectopic expression of angiogenesis can be deleterious to an organism. A number of pathological conditions are associated with the growth of extraneous blood vessels. These include, e.g., diabetic retinopathy, neovascular glaucoma, psoriasis, retrolental fibroplasias, angiofibroma, inflammation, etc. In addition, the increased blood supply associated with cancerous and neoplastic tissue, encourages growth, leading to rapid tumor enlargement and metastasis.

Because of the importance of angiogenesis in many physiological processes, its regulation has application in a vast arena of technologies and treatments. For instance, induction of neoangiogenesis has been used for the treatment of ischemic myocardial diseases, and other conditions (e.g., ischemic limb, stroke) produced by the lack of adequate blood supply. See, e.g., Rosengart et al., *Circulation*, 100(5):468-74, 1999. In growth new tissues from progenitor and stem cells, angiogenesis is one of the key processes necessary. Where vascularization is undesirable, such as for cancer and the mentioned pathological conditions, inhibition of angiogenesis has been used as a treatment therapy. See, e.g., U.S. Pat. No. 6,024,688 for treating neoplasms using angiogenesis inhibitors.

A number of different factors have been identified which stimulate angiogenesis, e.g., by activating normally quiescent endothelial cells, by acting as a chemoattractant to developing capillaries, by stimulating gene expression, etc. These factors include, e.g. fibroblast growth factors, such as FGF-1 and FGF-2, vascular endothelial growth factor (VEGF), platelet-derived endothelial cell growth factor (PD-ECGF), etc. Inhibition of angiogenesis has been achieved using drugs, such as TNP-470, monoclonal antibodies, antisense nucleic acids and proteins, such as angiostatin and endostatin. See, e.g., Battegay, *J. Mol. Med.*, 73, 333-346 (1995); Hanahan et al., *Cell*, 86, 353-364 (1996); Folkman, *N. Engl. J. Med.*, 333, 1757-1763 (1995).

Activity of a polynucleotide or gene in modulating or regulating angiogenesis can be determined according to any effective *in vivo* or *in vitro* methods. One useful model to study angiogenesis is based on the observation that, when a reconstituted basement membrane matrix, such as Matrigel[®], supplemented with growth factor (e.g., FGF-1), is injected

subcutaneously into a host animal, endothelial cells are recruited into the matrix, forming new blood vessels over a period of several days. See, e.g., Passaniti et al., *Lab. Invest.*, 67:519-528, 1992. By sampling the extract at different times, angiogenesis can be temporally dissected, permitting the identification of genes involved in all stages of angiogenesis, including, e.g., migration of endothelial cells into the matrix, commitment of endothelial cells to angiogenesis pathway, cell elongation and formation of sac-like spaces, and establishment of functional capillaries comprising connected, and linear structures containing red blood cells. To stabilize the growth factor and/or slow its release from the matrix, the growth factor can be bound to heparin or another stabilizing agent. The matrix can also be periodically re-infused with growth factor to enhance and extend the angiogenic process.

Other useful systems for studying angiogenesis, include, e.g., neovascularization of tumor explants (e.g., U.S. Pat. Nos. 5,192,744; 6,024,688), chicken chorioallantoic membrane (CAM) assay (e.g., Taylor and Folkman, *Nature*, 297:307-312, 1982; Eliceiri et al., *J. Cell Biol.*, 140, 1255-1263, 1998), bovine capillary endothelial (BCE) cell assay (e.g., U.S. Pat. No. 6,024,688; Polverini, P. J. et al., *Methods Enzymol.*, 198: 440-450, 1991), migration assays, HUVEC (human umbilical cord vascular endothelial cell) growth inhibition assay (e.g., U.S. Pat. No. 6,060,449).

The present invention relates to a human dehydrogenase (ANH401) which is expressed during angiogenesis. This gene was identified using a model system for angiogenesis. In this system, a Matrigel™ plug implant comprising FGF-1 is implanted subcutaneously into a host mouse. The initial bolus of FGF attracts endothelial cells into the implant, but does not result in new blood vessel formation. After about 10-15 days, the implant is re-infused with FGF-1. The FGF-1 stimulates the endothelial cells already present in the implant, initiating the process of angiogenesis. Tissue samples, removed at different intervals, can be analyzed to determine their gene expression patterns. ANH401 is differentially expressed in this system, e.g., being up-regulated under certain circumstances in the nascent blood vessels.

ANH401

ANH401 ("angiogenesis human gene 401"; also known as "ANH0401A") codes for a polypeptide containing 553 amino acids. It is up-regulated during angiogenesis, e.g.,

increasing in expression very early during the angiogenic process, and continuing at sustained levels throughout. The nucleotide and amino acid sequences of ANH401 are shown in SEQ ID NOS 1 and 2. It can have a Q or H at position 460. See, Fig. 1. It contains a PWWP-like domain at about amino acids 4-62 (e.g., involved in protein-protein interactions), an AT-hook-like domain (e.g., involved in DNA binding) at about amino acids 166-178, and a 6PGD ("6-phosphogluconate dehydrogenase"-like) domain at about amino acids 271-310. It contains the highly conserved NADP-binding domain, and other features of beta-hydroxyacid dehydrogenases. See, Njau et al., *Chemico-Biol. Inter.*, 130-132:785-791, 2001. In referring to ANH401 as a dehydrogenase, it is meant generally an activity involving NADP (or NAD) as a cofactor. The presence of a PWWP and AT-hook domains indicate that ANH401 may have nuclear localization, and/or nuclear function, such as DNA binding. See, e.g., Stec et al., *FEBS Letters*, 473:1-5, 2000. The dinucleotide cofactor (e.g., NAD or NADP binding at about amino acids 271-300) can also regulate the interaction of ANH401 with other proteins.

ANH401 is related to AF326966 and XM_048113. AF326966 shares about 98% amino acid sequence identity with ANH401 (calculated using the publicly available BLAST program). Compared with it, ANH401 has a six amino acid insertion from amino acid 303-308 in the 6PGD domain. See, Fig. 1. The absence of this sequence from AF326966 significantly diminishes its identity with, and function as, a 6PGD and associated other functions. XM_048113 is only a partial sequence, missing the PWWP, AT-hook, and a part of the 6PGD domains. See, Fig. 1. The mouse homolog of ANH401, BC006893, shares about 97% amino acid sequence identity with it and about 93% nucleotide sequence identity with it.

There are several UniGene clusters mapping to ANH401, Hs.331584 and Hs.87850. All or part of ANH401 is located in genomic DNA represented by GenBank ID: AC020663, BAC-ID: RP11-127I20, and Contig ID: NT_027178.3. The present invention relates to any isolated introns and exons that are present in the gene which, as discussed below. Intron and exon boundaries can be routinely determined, e.g., using the sequences disclosed herein.

Nucleic acids of the present invention map to chromosomal band 16p13.3. There are a number of different disorders which have been mapped to, or in close proximity to, this chromosome location. These include, e.g., congenital cataract with microphthalmia (CATM), hydroencephaly, and microcephaly. Nucleic acids of the present invention can be

used as linkage markers, diagnostic targets, therapeutic targets, for any of the mentioned disorders, as well as any disorders or genes mapping in proximity to it.

6-phosphogluconate dehydrogenase (6PGD), along with glucose-6-phosphate dehydrogenase (G6PD), are major suppliers of NADPH to the cell. NADPH is a powerful reductant that is utilized in a variety of metabolic and regulatory pathways. Much of the NADPH manufactured in a cell is produced by the pentose phosphate pathway. In this pathway, G6PD first catalyzes the formation of 6-phosphoglucolactone from glucose-6-phosphate, yielding 1 mole of NADPH from NADP^+ . The 6-phosphoglucolactone is rapidly converted into 6-phosphogluconic acid (6-PGA) which is substrate for 6PGD. The subsequent conversion of 6-PGA into ribulose phosphate by 6PGD is accompanied by the production of NADPH from NADP^+ . See, e.g., Martini and Ursini, *BioEssays*, 18:631-637, 1996.

G6PD deficiency is one of the most common enzyme disorders found in humans. Deficiency in 6PGD is much less common than G6PD. Although most patients with G6PD deficiency are asymptomatic, certain drugs and environmental exposures can lead to hemolytic anemia in patients with defective G6PD levels, including such drugs as sulfonamides, aspirin, non-steroidal anti-inflammatory drugs (NSAIDs), nitrofurantoin, quinidine, and quinine. Due to the prevalence of G6PD deficiency in the population, screening for it can be important, especially when a patient is to be treated with a drug that is known to have adverse effects in patients having the deficiency. Various commercial assays for it are available, many of which are spectrophotometric, based on the absorption of NADPH at 340 nm.

Assays for G6PD and 6PGD activities can be routinely performed, e.g., on erythrocytes, solid tissues, etc. As indicated above, such assays are routine and make use of the spectrophotometric change when NADPH is produced. A number of different methods can be used to dissociate the two activities from each other, e.g., by providing enzyme specific substrates (see above), competitors, etc. Purified enzyme preparations can be used as controls and to determine reaction parameters. See, e.g., OxisResearch™, Bioxytech® G6PD/6PGD-340™.

30

Nucleic acids

A mammalian polynucleotide, or fragment thereof, of the present invention is a polynucleotide having a nucleotide sequence obtainable from a natural source. When the species name is used, e.g., human ANH401, it indicates that the polynucleotide or polypeptide is obtainable from a natural source. It therefore includes naturally-occurring normal, naturally-occurring mutant, and naturally-occurring polymorphic alleles (e.g., SNPs), differentially-spliced transcripts, splice-variants, etc. By the term "naturally-occurring," it is meant that the polynucleotide is obtainable from a natural source, e.g., animal tissue and cells, body fluids, tissue culture cells, forensic samples. Natural sources include, e.g., living cells obtained from tissues and whole organisms, tumors, cultured cell lines, including primary and immortalized cell lines. Naturally-occurring mutations can include deletions (e.g., a truncated amino- or carboxy-terminus), substitutions, inversions, or additions of nucleotide sequence. These genes can be detected and isolated by polynucleotide hybridization according to methods which one skilled in the art would know, e.g., as discussed below.

A polynucleotide according to the present invention can be obtained from a variety of different sources. It can be obtained from DNA or RNA, such as polyadenylated mRNA or total RNA, e.g., isolated from tissues, cells, or whole organism. The polynucleotide can be obtained directly from DNA or RNA, from a cDNA library, from a genomic library, etc. The polynucleotide can be obtained from a cell or tissue (e.g., from an embryonic or adult tissues) at a particular stage of development, having a desired genotype, phenotype, disease status, etc. A polynucleotide which "codes without interruption" refers to a polynucleotide having a continuous open reading frame ("ORF") as compared to an ORF which is interrupted by introns or other noncoding sequences.

Polynucleotides and polypeptides (including any part of ANH401) can be excluded as compositions from the present invention if, e.g., listed in a publicly available databases on the day this application was filed and/or disclosed in a patent application having an earlier filing or priority date than this application and/or conceived and/or reduced to practice earlier than a polynucleotide in this application.

As described herein, the phrase "an isolated polynucleotide which is SEQ ID NO," or "an isolated polynucleotide which is selected from SEQ ID NO," refers to an isolated nucleic acid molecule from which the recited sequence was derived (e.g., a cDNA derived from

mRNA; cDNA derived from genomic DNA). Because of sequencing errors, typographical errors, etc., the actual naturally-occurring sequence may differ from a SEQ ID listed herein. Thus, the phrase indicates the specific molecule from which the sequence was derived, rather than a molecule having that exact recited nucleotide sequence, analogously to how a culture depository number refers to a specific cloned fragment in a cryotube.

As explained in more detail below, a polynucleotide sequence of the invention can contain the complete sequence as shown in SEQ ID NO 1, degenerate sequences thereof, anti-sense, muteins thereof, genes comprising said sequences, full-length cDNAs comprising said sequences, complete genomic sequences, fragments thereof, homologs, primers, nucleic acid molecules which hybridize thereto, derivatives thereof, etc.

Genomic

The present invention also relates genomic DNA from which the polynucleotides of the present invention can be derived. A genomic DNA coding for a human, mouse, or other mammalian polynucleotide, can be obtained routinely, for example, by screening a genomic library (e.g., a YAC library) with a polynucleotide of the present invention, or by searching nucleotide databases, such as GenBank and EMBL, for matches. Promoter and other regulatory regions (including both 5' and 3' regions, as well introns) can be identified upstream of coding and expressed RNAs, and assayed routinely for activity, e.g., by joining to a reporter gene (e.g., CAT, GFP, alkaline phosphatase, luciferase, galactosidase). A promoter obtained from ANH401 can be used, e.g., in gene therapy to obtain tissue-specific expression of a heterologous gene (e.g., coding for a therapeutic product or cytotoxin). 5' and/or 3' sequences can also be used to modulate stability of a nucleic acid, regulate its translation and/or transcription, etc.

Constructs

A polynucleotide of the present invention can comprise additional polynucleotide sequences, e.g., sequences to enhance expression, detection, uptake, cataloging, tagging, etc. A polynucleotide can include only coding sequence; a coding sequence and additional non-naturally occurring or heterologous coding sequence (e.g., sequences coding for leader, signal, secretory, targeting, enzymatic, fluorescent, antibiotic resistance, and other functional

or diagnostic peptides); coding sequences and non-coding sequences, e.g., untranslated sequences at either a 5' or 3' end, or dispersed in the coding sequence, e.g., introns.

A polynucleotide according to the present invention also can comprise an expression control sequence operably linked to a polynucleotide as described above. The phrase

5 "expression control sequence" means a polynucleotide sequence that regulates expression of a polypeptide coded for by a polynucleotide to which it is functionally ("operably") linked. Expression can be regulated at the level of the mRNA or polypeptide. Thus, the expression control sequence includes mRNA-related elements and protein-related elements. Such elements include promoters, enhancers (viral or cellular), ribosome binding sequences, transcriptional terminators, etc. An expression control sequence is operably linked to a nucleotide coding sequence when the expression control sequence is positioned in such a manner to effect or achieve expression of the coding sequence. For example, when a promoter is operably linked 5' to a coding sequence, expression of the coding sequence is driven by the promoter. Expression control sequences can include an initiation codon and additional nucleotides to place a partial nucleotide sequence of the present invention in-frame in order to produce a polypeptide (e.g., pET vectors from Promega have been designed to permit a molecule to be inserted into all three reading frames to identify the one that results in polypeptide expression). Expression control sequences can be heterologous or endogenous to the normal gene.

20 A polynucleotide of the present invention can also comprise nucleic acid vector sequences, e.g., for cloning, expression, amplification, selection, etc. Any effective vector can be used. A vector is, e.g., a polynucleotide molecule which can replicate autonomously in a host cell, e.g., containing an origin of replication. Vectors can be useful to perform manipulations, to propagate, and/or obtain large quantities of the recombinant molecule in a desired host. A skilled worker can select a vector depending on the purpose desired, e.g., to propagate the recombinant molecule in bacteria, yeast, insect, or mammalian cells. The following vectors are provided by way of example. Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBS, pD10, Phagescript, phiX174, pBK Phagemid, pNH8A, pNH16a, pNH18Z, pNH46A (Stratagene); Bluescript KS+II (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR54 0, pRIT5 (Pharmacia). Eukaryotic: PWLNEO, pSV2CAT, pOG44, pXT1, pSG 30 (Stratagene), pSVK3, PBPV, PMSG, pSVL (Pharmacia), pCR2.1/TOPO, pCRII/TOPO,

pCR4/TOPO, pTrcHisB, pCMV6-XL4, etc. However, any other vector, e.g., plasmids, viruses, or parts thereof, may be used as long as they are replicable and viable in the desired host. The vector can also comprise sequences which enable it to replicate in the host whose genome is to be modified.

5

Hybridization

Polynucleotide hybridization, as discussed in more detail below, is useful in a variety of applications, including, in gene detection methods, for identifying mutations, for making mutations, to identify homologs in the same and different species, to identify
10 related members of the same gene family, in diagnostic and prognostic assays, in therapeutic applications (e.g., where an antisense polynucleotide is used to inhibit expression), etc.

The ability of two single-stranded polynucleotide preparations to hybridize together is a measure of their nucleotide sequence complementarity, e.g., base-pairing between
15 nucleotides, such as A-T, G-C, etc. The invention thus also relates to polynucleotides, and their complements, which hybridize to a polynucleotide comprising a nucleotide sequence as set forth in SEQ ID NO 1 and genomic sequences thereof. A nucleotide sequence hybridizing to the latter sequence will have a complementary polynucleotide strand, or act as a template for one in the presence of a polymerase (i.e., an appropriate
20 polynucleotide synthesizing enzyme). The present invention includes both strands of polynucleotide, e.g., a sense strand and an anti-sense strand.

Hybridization conditions can be chosen to select polynucleotides which have a desired amount of nucleotide complementarity with the nucleotide sequences set forth in SEQ ID NO 1 and genomic sequences thereof. A polynucleotide capable of hybridizing
25 to such sequence, preferably, possesses, e.g., about 70%, 75%, 80%, 85%, 87%, 90%, 92%, 95%, 97%, 99%, or 100% complementarity, between the sequences. The present invention particularly relates to polynucleotide sequences which hybridize to the nucleotide sequences set forth in SEQ ID NO 1 or genomic sequences thereof, under low or high stringency conditions. These conditions can be used, e.g., to select corresponding
30 homologs in non-human species.

Polynucleotides which hybridize to polynucleotides of the present invention can be selected in various ways. Filter-type blots (i.e., matrices containing polynucleotide, such as nitrocellulose), glass chips, and other matrices and substrates comprising polynucleotides (short or long) of interest, can be incubated in a prehybridization solution (e.g., 6X SSC, 0.5% SDS, 100 µg/ml denatured salmon sperm DNA, 5X Denhardt's solution, and 50% formamide), at 22-68°C, overnight, and then hybridized with a detectable polynucleotide probe under conditions appropriate to achieve the desired stringency. In general, when high homology or sequence identity is desired, a high temperature can be used (e.g., 65 °C). As the homology drops, lower washing temperatures are used. For salt concentrations, the lower the salt concentration, the higher the stringency. The length of the probe is another consideration. Very short probes (e.g., less than 100 base pairs) are washed at lower temperatures, even if the homology is high. With short probes, formamide can be omitted. See, e.g., *Current Protocols in Molecular Biology*, Chapter 6, Screening of Recombinant Libraries; Sambrook et al., *Molecular Cloning*, 1989, Chapter 9.

For instance, high stringency conditions can be achieved by incubating the blot overnight (e.g., at least 12 hours) with a long polynucleotide probe in a hybridization solution containing, e.g., about 5X SSC, 0.5% SDS, 100 µg/ml denatured salmon sperm DNA and 50% formamide, at 42°C. Blots can be washed at high stringency conditions that allow, e.g., for less than 5% bp mismatch (e.g., wash twice in 0.1% SSC and 0.1% SDS for 30 min at 65°C), i.e., selecting sequences having 95% or greater sequence identity.

Other non-limiting examples of high stringency conditions includes a final wash at 65°C in aqueous buffer containing 30 mM NaCl and 0.5% SDS. Another example of high stringent conditions is hybridization in 7% SDS, 0.5 M NaPO₄, pH 7, 1 mM EDTA at 50°C, e.g., overnight, followed by one or more washes with a 1% SDS solution at 42°C.

Whereas high stringency washes can allow for less than 5% mismatch, reduced or low stringency conditions can permit up to 20% nucleotide mismatch. Hybridization at low stringency can be accomplished as above, but using lower formamide conditions, lower temperatures and/or lower salt concentrations, as well as longer periods of incubation time.

Hybridization can also be based on a calculation of melting temperature (T_m) of the hybrid formed between the probe and its target, as described in Sambrook et al.. Generally, the temperature T_m at which a short oligonucleotide (containing 18 nucleotides

or fewer) will melt from its target sequence is given by the following equation: $T_m =$
(number of A's and T's) \times 2°C + (number of C's and G's) \times 4°C . For longer molecules,
 $T_m = 81.5 + 16.6 \log_{10}[\text{Na}^+] + 0.41(\%GC) - 600/N$ where $[\text{Na}^+]$ is the molar
concentration of sodium ions, %GC is the percentage of GC base pairs in the probe, and
5 N is the length. Hybridization can be carried out at several degrees below this
temperature to ensure that the probe and target can hybridize. Mismatches can be allowed
for by lowering the temperature even further.

Stringent conditions can be selected to isolate sequences, and their complements,
which have, e.g., at least about 90%, 95%, or 97%, nucleotide complementarity between
10 the probe (e.g., a short polynucleotide of SEQ ID NO 1 or genomic sequences thereof)
and a target polynucleotide.

Other homologs of polynucleotides of the present invention can be obtained from
mammalian and non-mammalian sources according to various methods. For example,
hybridization with a polynucleotide can be employed to select homologs, e.g., as described in
15 Sambrook et al., *Molecular Cloning*, Chapter 11, 1989. Such homologs can have varying
amounts of nucleotide and amino acid sequence identity and similarity to such
polynucleotides of the present invention. Mammalian organisms include, e.g., mice, rats,
monkeys, pigs, cows, etc. Non-mammalian organisms include, e.g., vertebrates,
invertebrates, zebra fish, chicken, *Drosophila*, *C. elegans*, *Xenopus*, yeast such as *S. pombe*,
20 *S. cerevisiae*, roundworms, prokaryotes, plants, *Arabidopsis*, *artemia*, viruses, etc.

Alignment

Alignments can be accomplished by using any effective algorithm. For pairwise
alignments of DNA sequences, the methods described by Wilbur-Lipman (e.g., Wilbur and
25 Lipman, *Proc. Natl. Acad. Sci.*, 80:726-730, 1983) or Martinez/Needleman-Wunsch (e.g.,
Martinez, *Nucleic Acid Res.*, 11:4629-4634, 1983) can be used. For instance, if the
Martinez/Needleman-Wunsch DNA alignment is applied, the minimum match can be set at
9, gap penalty at 1.10, and gap length penalty at 0.33. The results can be calculated as a
similarity index, equal to the sum of the matching residues divided by the sum of all residues
30 and gap characters, and then multiplied by 100 to express as a percent. Similarity index for
related genes at the nucleotide level in accordance with the present invention can be greater

than 70%, 80%, 85%, 90%, 95%, 99%, or more. Pairs of protein sequences can be aligned by the Lipman-Pearson method (e.g., Lipman and Pearson, *Science*, 227:1435-1441, 1985) with k-tuple set at 2, gap penalty set at 4, and gap length penalty set at 12. Results can be expressed as percent similarity index, where related genes at the amino acid level in accordance with the present invention can be greater than 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99%, or more. Various commercial and free sources of alignment programs are available, e.g., MegAlign by DNA Star, BLAST (National Center for Biotechnology Information), BCM (Baylor College of Medicine) Launcher, etc. BLAST can be used to calculate amino acid sequence identity, amino acid sequence homology, and nucleotide sequence identity. These calculations can be made along the entire length of each of the target sequences which are to be compared.

Percent sequence identity can also be determined by other conventional methods, e.g., as described in Altschul et al., *Bull. Math. Bio.* 48: 603-616, 1986 and Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915-10919, 1992.

Specific polynucleotide probes

A polynucleotide of the present invention can comprise any continuous nucleotide sequence of SEQ ID NO 1, sequences which share sequence identity thereto, or complements thereof. The term "probe" refers to any substance that can be used to detect, identify, isolate, etc., another substance. A polynucleotide probe is comprised of nucleic acid can be used to detect, identify, etc., other nucleic acids, such as DNA and RNA.

These polynucleotides can be of any desired size that is effective to achieve the specificity desired. For example, a probe can be from about 7 or 8 nucleotides to several thousand nucleotides, depending upon its use and purpose. For instance, a probe used as a primer PCR can be shorter than a probe used in an ordered array of polynucleotide probes. Probe sizes vary, and the invention is not limited in any way by their size, e.g., probes can be from about 7-2000 nucleotides, 7-1000, 8-700, 8-600, 8-500, 8-400, 8-300, 8-150, 8-100, 8-75, 7-50, 10-25, 14-16, at least about 8, at least about 10, at least about 15, at least about 25, etc. The polynucleotides can have non-naturally-occurring nucleotides, e.g., inosine, AZT, 3TC, etc. The polynucleotides can have 100% sequence identity or complementarity to a sequence of SEQ ID NO 1, or it can have mismatches or nucleotide substitutions, e.g., 1, 2, 3,

4, or 5 substitutions. The probes can be single-stranded or double-stranded. Polynucleotides can code for, e.g., amino acids 303-308, 271-300, 271-282, 283-300, 271-308 of SEQ ID NO 2.

In accordance with the present invention, a polynucleotide can be present in a kit, where the kit includes, e.g., one or more polynucleotides, a desired buffer (e.g., phosphate, tris, etc.), detection compositions, RNA or cDNA from different tissues to be used as controls, libraries, etc. The polynucleotide can be labeled or unlabeled, with radioactive or non-radioactive labels as known in the art. Kits can comprise one or more pairs of polynucleotides for amplifying nucleic acids specific for ANH401, e.g., comprising a forward and reverse primer effective in PCR. These include both sense and anti-sense orientations. For instance, in PCR-based methods (such as RT-PCR), a pair of primers are typically used, one having a sense sequence and the other having an antisense sequence.

Another aspect of the present invention is a nucleotide sequence that is specific to, or for, a selective polynucleotide. The phrases "specific for" or "specific to" a polynucleotide have a functional meaning that the polynucleotide can be used to identify the presence of one or more target genes in a sample and distinguish them from non-target genes. It is specific in the sense that it can be used to detect polynucleotides above background noise ("non-specific binding"). A specific sequence is a defined order of nucleotides (or amino acid sequences, if it is a polypeptide sequence) which occurs in the polynucleotide, e.g., in the nucleotide sequences of SEQ ID NO 1, and which is characteristic of that target sequence, and substantially no non-target sequences. A probe or mixture of probes can comprise a sequence or sequences that are specific to a plurality of target sequences, e.g., where the sequence is a consensus sequence, a functional domain, etc., e.g., capable of recognizing a family of related genes. Such sequences can be used as probes in any of the methods described herein or incorporated by reference. Both sense and antisense nucleotide sequences are included. A specific polynucleotide according to the present invention can be determined routinely.

A polynucleotide comprising a specific sequence can be used as a hybridization probe to identify the presence of, e.g., human or mouse polynucleotide, in a sample comprising a mixture of polynucleotides, e.g., on a Northern blot. Hybridization can be performed under high stringent conditions (see, above) to select polynucleotides (and their complements which

can contain the coding sequence) having at least 90%, 95%, 99%, etc., identity (i.e., complementarity) to the probe, but less stringent conditions can also be used. A specific polynucleotide sequence can also be fused in-frame, at either its 5' or 3' end, to various nucleotide sequences as mentioned throughout the patent, including coding sequences for enzymes, detectable markers, GFP, etc, expression control sequences, etc.

A polynucleotide probe, especially one that is specific to a polynucleotide of the present invention, can be used in gene detection and hybridization methods as already described. In one embodiment, a specific polynucleotide probe can be used to detect whether a particular tissue or cell-type is present in a target sample. To carry out such a method, a selective polynucleotide can be chosen which is characteristic of the desired target tissue. Such polynucleotide is preferably chosen so that it is expressed or displayed in the target tissue, but not in other tissues which are present in the sample. For instance, if detection of vascular is desired, it may not matter whether the selective polynucleotide is expressed in other tissues, as long as it is not expressed in cells normally present in blood, e.g., peripheral blood mononuclear cells. Starting from the selective polynucleotide, a specific polynucleotide probe can be designed which hybridizes (if hybridization is the basis of the assay) under the hybridization conditions to the selective polynucleotide, whereby the presence of the selective polynucleotide can be determined.

Probes which are specific for polynucleotides of the present invention can also be prepared using involve transcription-based systems, e.g., incorporating an RNA polymerase promoter into a selective polynucleotide of the present invention, and then transcribing anti-sense RNA using the polynucleotide as a template. See, e.g., U.S. Pat. No. 5,545,522.

Polynucleotide composition

A polynucleotide according to the present invention can comprise, e.g., DNA, RNA, synthetic polynucleotide, peptide polynucleotide, modified nucleotides, dsDNA, ssDNA, ssRNA, dsRNA, and mixtures thereof. A polynucleotide can be single- or double-stranded, triplex, DNA:RNA, duplexes, comprise hairpins, and other secondary structures, etc. Nucleotides comprising a polynucleotide can be joined via various known linkages, e.g., ester, sulfamate, sulfamide, phosphorothioate, phosphoramidate, methylphosphonate, carbamate, etc., depending on the desired purpose, e.g., resistance to nucleases, such as

RNAse H, improved in vivo stability, etc. See, e.g., U.S. Pat. No. 5,378,825. Any desired nucleotide or nucleotide analog can be incorporated, e.g., 6-mercaptopguanine, 8-oxo-guanine, etc.

Various modifications can be made to the polynucleotides, such as attaching
5 detectable markers (avidin, biotin, radioactive elements, fluorescent tags and dyes, energy transfer labels, energy-emitting labels, binding partners, etc.) or moieties which improve hybridization, detection, and/or stability. The polynucleotides can also be attached to solid supports, e.g., nitrocellulose, magnetic or paramagnetic microspheres (e.g., as described in U.S. Pat. No. 5,411,863; U.S. Pat. No. 5,543,289; for instance, comprising ferromagnetic,
10 supermagnetic, paramagnetic, superparamagnetic, iron oxide and polysaccharide), nylon, agarose, diazotized cellulose, latex solid microspheres, polyacrylamides, etc., according to a desired method. See, e.g., U.S. Pat. Nos. 5,470,967, 5,476,925, and 5,478,893.

Polynucleotide according to the present invention can be labeled according to any desired method. The polynucleotide can be labeled using radioactive tracers such as ^{32}P , ^{35}S ,
15 ^3H , or ^{14}C , to mention some commonly used tracers. The radioactive labeling can be carried out according to any method, such as, for example, terminal labeling at the 3' or 5' end using a radiolabeled nucleotide, polynucleotide kinase (with or without dephosphorylation with a phosphatase) or a ligase (depending on the end to be labeled). A non-radioactive labeling can also be used, combining a polynucleotide of the present invention with residues having
20 immunological properties (antigens, haptens), a specific affinity for certain reagents (ligands), properties enabling detectable enzyme reactions to be completed (enzymes or coenzymes, enzyme substrates, or other substances involved in an enzymatic reaction), or characteristic physical properties, such as fluorescence or the emission or absorption of light at a desired wavelength, etc.

25 Nucleic acid detection methods

Another aspect of the present invention relates to methods and processes for detecting ANH401. Detection methods have a variety of applications, including for diagnostic, prognostic, forensic, and research applications. To accomplish gene detection, a
30 polynucleotide in accordance with the present invention can be used as a "probe." The term "probe" or "polynucleotide probe" has its customary meaning in the art, e.g., a

polynucleotide which is effective to identify (e.g., by hybridization), when used in an appropriate process, the presence of a target polynucleotide to which it is designed. Identification can involve simply determining presence or absence, or it can be quantitative, e.g., in assessing amounts of a gene or gene transcript present in a sample. Probes can be useful in a variety of ways, such as for diagnostic purposes, to identify homologs, and to detect, quantitate, or isolate a polynucleotide of the present invention in a test sample.

Assays can be utilized which permit quantification and/or presence/absence detection of a target nucleic acid in a sample. Assays can be performed at the single-cell level, or in a sample comprising many cells, where the assay is "averaging" expression over the entire collection of cells and tissue present in the sample. Any suitable assay format can be used, including, but not limited to, e.g., Southern blot analysis, Northern blot analysis, polymerase chain reaction ("PCR") (e.g., Saiki et al., *Science*, 241:53, 1988; U.S. Pat. Nos. 4,683,195, 4,683,202, and 6,040,166; *PCR Protocols: A Guide to Methods and Applications*, Innis et al., eds., Academic Press, New York, 1990), reverse transcriptase polymerase chain reaction ("RT-PCR"), anchored PCR, rapid amplification of cDNA ends ("RACE") (e.g., Schaefer in *Gene Cloning and Analysis: Current Innovations*, Pages 99-115, 1997), ligase chain reaction ("LCR") (EP 320 308), one-sided PCR (Ohara et al., *Proc. Natl. Acad. Sci.*, 86:5673-5677, 1989), indexing methods (e.g., U.S. Pat. No. 5,508,169), *in situ* hybridization, differential display (e.g., Liang et al., *Nucl. Acid. Res.*, 21:3269-3275, 1993; U.S. Pat. Nos. 5,262,311, 5,599,672 and 5,965,409; WO97/18454; Prashar and Weissman, *Proc. Natl. Acad. Sci.*, 93:659-663, and U.S. Pat. Nos. 6,010,850 and 5,712,126; Welsh et al., *Nucleic Acid Res.*, 20:4965-4970, 1992, and U.S. Pat. No. 5,487,985) and other RNA fingerprinting techniques, nucleic acid sequence based amplification ("NASBA") and other transcription based amplification systems (e.g., U.S. Pat. Nos. 5,409,818 and 5,554,527; WO 88/10315), polynucleotide arrays (e.g., U.S. Pat. Nos. 5,143,854, 5,424,186; 5,700,637, 5,874,219, and 6,054,270; PCT WO 92/10092; PCT WO 90/15070), Qbeta Replicase (PCT/US87/00880), Strand Displacement Amplification ("SDA"), Repair Chain Reaction ("RCR"), nuclease protection assays, subtraction-based methods, Rapid-Scan™, etc. Additional useful methods include, but are not limited to, e.g., template-based amplification methods, competitive PCR (e.g., U.S. Pat. No. 5,747,251), redox-based assays (e.g., U.S. Pat. No. 5,871,918), Taqman-based assays (e.g., Holland et al., *Proc. Natl. Acad. Sci.*, 88:7276-7280, 1991; U.S. Pat. Nos.

5,210,015 and 5,994,063), real-time fluorescence-based monitoring (e.g., U.S. Pat. 5,928,907), molecular energy transfer labels (e.g., U.S. Pat. Nos. 5,348,853, 5,532,129, 5,565,322, 6,030,787, and 6,117,635; Tyagi and Kramer, *Nature Biotech.*, 14:303-309, 1996). Any method suitable for single cell analysis of gene or protein expression can be
5 used, including in situ hybridization, immunocytochemistry, MACS, FACS, flow cytometry, etc. For single cell assays, expression products can be measured using antibodies, PCR, or other types of nucleic acid amplification (e.g., Brady et al., *Methods Mol. & Cell. Biol.* 2, 17-25, 1990; Eberwine et al., 1992, *Proc. Natl. Acad. Sci.*, 89, 3010-3014, 1992; U.S. Pat. No. 5,723,290). These and other methods can be carried out conventionally, e.g., as described in
10 the mentioned publications.

Many of such methods may require that the polynucleotide is labeled, or comprises a particular nucleotide type useful for detection. The present invention includes such modified polynucleotides that are necessary to carry out such methods. Thus, polynucleotides can be DNA, RNA, DNA:RNA hybrids, PNA, etc., and can comprise any modification or
15 substituent which is effective to achieve detection.

Detection can be desirable for a variety of different purposes, including research, diagnostic, prognostic, and forensic. For diagnostic purposes, it may be desirable to identify the presence or quantity of a polynucleotide sequence in a sample, where the sample is obtained from tissue, cells, body fluids, etc. In a preferred method as described in more
20 detail below, the present invention relates to a method of detecting a polynucleotide comprising, contacting a target polynucleotide in a test sample with a polynucleotide probe under conditions effective to achieve hybridization between the target and probe; and detecting hybridization.

Any test sample in which it is desired to identify a polynucleotide or polypeptide
25 thereof can be used, including, e.g., blood, urine, saliva, stool (for extracting nucleic acid, see, e.g., U.S. Pat. No. 6,177,251), swabs comprising tissue, biopsied tissue, tissue sections, cultured cells, etc.

Detection can be accomplished in combination with polynucleotide probes for other genes, e.g., genes which are expressed in other disease states, tissues, cells, such as brain,
30 heart, kidney, spleen, thymus, liver, stomach, small intestine, colon, muscle, lung, testis, placenta, pituitary, thyroid, skin, adrenal gland, pancreas, salivary gland, uterus, ovary,

prostate gland, peripheral blood cells (T-cells, lymphocytes, etc.), embryo, normal breast fat, adult and embryonic stem cells, specific cell-types, such as endothelial, epithelial, myocytes, adipose, luminal epithelial, basoepithelial, myoepithelial, stromal cells, etc.

Polynucleotides can be used in wide range of methods and compositions, including
5 for detecting, diagnosing, staging, grading, assessing, prognosticating, etc. diseases and disorders associated with ANH401, for monitoring or assessing therapeutic and/or preventative measures, in ordered arrays, etc. Any method of detecting genes and polynucleotides of SEQ ID NO 1 can be used; certainly, the present invention is not to be limited how such methods are implemented.

10 Along these lines, the present invention relates to methods of detecting ANH401 in a sample comprising nucleic acid. Such methods can comprise one or more the following steps in any effective order, e.g., contacting said sample with a polynucleotide probe under conditions effective for said probe to hybridize specifically to nucleic acid in said sample, and detecting the presence or absence of probe hybridized to nucleic acid in said sample,
15 wherein said probe is a polynucleotide which is SEQ ID NO 1, a polynucleotide having, e.g., about 70%, 80%, 85%, 90%, 95%, 99%, or more sequence identity thereto, effective or specific fragments thereof, or complements thereto. The detection method can be applied to any sample, e.g., cultured primary, secondary, or established cell lines, tissue biopsy, blood, urine, stool, cerebral spinal fluid, and other bodily fluids, for any purpose.

20 Contacting the sample with probe can be carried out by any effective means in any effective environment. It can be accomplished in a solid, liquid, frozen, gaseous, amorphous, solidified, coagulated, colloid, etc., mixtures thereof, matrix. For instance, a probe in an aqueous medium can be contacted with a sample which is also in an aqueous medium, or which is affixed to a solid matrix, or vice-versa.

25 Generally, as used throughout the specification, the term "effective conditions" means, e.g., the particular milieu in which the desired effect is achieved. Such a milieu, includes, e.g., appropriate buffers, oxidizing agents, reducing agents, pH, co-factors, temperature, ion concentrations, suitable age and/or stage of cell (such as, in particular part of the cell cycle, or at a particular stage where particular genes are being expressed) where cells
30 are being used, culture conditions (including substrate, oxygen, carbon dioxide, etc.). When hybridization is the chosen means of achieving detection, the probe and sample can be

combined such that the resulting conditions are functional for said probe to hybridize specifically to nucleic acid in said sample.

The phrase "hybridize specifically" indicates that the hybridization between single-stranded polynucleotides is based on nucleotide sequence complementarity. The effective conditions are selected such that the probe hybridizes to a preselected and/or definite target nucleic acid in the sample. For instance, if detection of a polynucleotide set forth in SEQ ID NO 1 is desired, a probe can be selected which can hybridize to such target gene under high stringent conditions, without significant hybridization to other genes in the sample. To detect homologs of a polynucleotide set forth in SEQ ID NO 1, the effective hybridization conditions can be less stringent, and/or the probe can comprise codon degeneracy, such that a homolog is detected in the sample.

As already mentioned, the methods can be carried out by any effective process, e.g., by Northern blot analysis, polymerase chain reaction (PCR), reverse transcriptase PCR, RACE PCR, *in situ* hybridization, etc., as indicated above. When PCR based techniques are used, two or more probes are generally used. One probe can be specific for a defined sequence which is characteristic of a selective polynucleotide, but the other probe can be specific for the selective polynucleotide, or specific for a more general sequence, e.g., a sequence such as polyA which is characteristic of mRNA, a sequence which is specific for a promoter, ribosome binding site, or other transcriptional features, a consensus sequence (e.g., representing a functional domain). For the former aspects, 5' and 3' probes (e.g., polyA, Kozak, etc.) are preferred which are capable of specifically hybridizing to the ends of transcripts. When PCR is utilized, the probes can also be referred to as "primers" in that they can prime a DNA polymerase reaction.

In addition to testing for the presence or absence of polynucleotides, the present invention also relates to determining the amounts at which polynucleotides of the present invention are expressed in sample and determining the differential expression of such polynucleotides in samples.. Such methods can involve substantially the same steps as described above for presence/absence detection, e.g., contacting with probe, hybridizing, and detecting hybridized probe, but using more quantitative methods and/or comparisons to standards.

The amount of hybridization between the probe and target can be determined by any suitable methods, e.g., PCR, RT-PCR, RACE PCR, Northern blot, polynucleotide microarrays, Rapid-Scan, etc., and includes both quantitative and qualitative measurements. For further details, see the hybridization methods described above and below. Determining by such hybridization whether the target is differentially expressed (e.g., up-regulated or down-regulated) in the sample can also be accomplished by any effective means. For instance, the target's expression pattern in the sample can be compared to its pattern in a known standard, such as in a normal tissue, or it can be compared to another gene in the same sample. When a second sample is utilized for the comparison, it can be a sample of normal tissue that is known not to contain diseased cells. The comparison can be performed on samples which contain the same amount of RNA (such as polyadenylated RNA or total RNA), or, on RNA extracted from the same amounts of starting tissue. Such a second sample can also be referred to as a control or standard. Hybridization can also be compared to a second target in the same tissue sample. Experiments can be performed that determine a ratio between the target nucleic acid and a second nucleic acid (a standard or control), e.g., in a normal tissue. When the ratio between the target and control are substantially the same in a normal and sample, the sample is determined or diagnosed not to contain cells. However, if the ratio is different between the normal and sample tissues, the sample is determined to contain cancer cells. The approaches can be combined, and one or more second samples, or second targets can be used. Any second target nucleic acid can be used as a comparison, including "housekeeping" genes, such as beta-actin, alcohol dehydrogenase, or any other gene whose expression does not vary depending upon the disease status of the cell.

Methods of identifying polymorphisms, mutations, etc., of ANH401

Polynucleotides of the present invention can also be utilized to identify mutant alleles, SNPs, gene rearrangements and modifications, and other polymorphisms of the wild-type gene. Mutant alleles, polymorphisms, SNPs, etc., can be identified and isolated from cancers that are known, or suspected to have, a genetic component. Identification of such genes can be carried out routinely (see, above for more guidance), e.g., using PCR, hybridization techniques, direct sequencing, mismatch reactions (see, e.g., above), RFLP analysis, SSCP (e.g., Orita et al., *Proc. Natl. Acad. Sci.*, 86:2766, 1992), etc., where a polynucleotide having

a sequence selected from SEQ ID NO 1 is used as a probe. The selected mutant alleles, SNPs, polymorphisms, etc., can be used diagnostically to determine whether a subject has, or is susceptible to a disorder associated with ANH401, as well as to design therapies and predict the outcome of the disorder. Methods involve, e.g., diagnosing a disorder associated with ANH401 or determining susceptibility to a disorder, comprising, detecting the presence of a mutation in a gene represented by a polynucleotide selected from SEQ ID NO 1. The detecting can be carried out by any effective method, e.g., obtaining cells from a subject, determining the gene sequence or structure of a target gene (using, e.g., mRNA, cDNA, genomic DNA, etc), comparing the sequence or structure of the target gene to the structure of the normal gene, whereby a difference in sequence or structure indicates a mutation in the gene in the subject. Polynucleotides can also be used to test for mutations, SNPs, polymorphisms, etc., e.g., using mismatch DNA repair technology as described in U.S. Pat. No. 5,683,877; U.S. Pat. No. 5,656,430; Wu et al., *Proc. Natl. Acad. Sci.*, 89:8779-8783, 1992.

The present invention also relates to methods of detecting polymorphisms in ANH401, comprising, e.g., comparing the structure of: genomic DNA comprising all or part of ANH401, mRNA comprising all or part of ANH401, cDNA comprising all or part of ANH401, or a polypeptide comprising all or part of ANH401, with the structure of ANH401 set forth in SEQ ID NO 1. The methods can be carried out on a sample from any source, e.g., cells, tissues, body fluids, blood, urine, stool, hair, egg, sperm, etc.

These methods can be implemented in many different ways. For example, "comparing the structure" steps include, but are not limited to, comparing restriction maps, nucleotide sequences, amino acid sequences, RFLPs, DNase sites, DNA methylation fingerprints (e.g., U.S. Pat. No. 6,214,556), protein cleavage sites, molecular weights, electrophoretic mobilities, charges, ion mobility, etc., between a standard ANH401 and a test ANH401. The term "structure" can refer to any physical characteristics or configurations which can be used to distinguish between nucleic acids and polypeptides. The methods and instruments used to accomplish the comparing step depends upon the physical characteristics which are to be compared. Thus, various techniques are contemplated, including, e.g., sequencing machines (both amino acid and polynucleotide), electrophoresis, mass spectrometer (U.S. Pat. Nos. 6,093,541,

6,002,127), liquid chromatography, HPLC, etc.

To carry out such methods, "all or part" of the gene or polypeptide can be compared. For example, if nucleotide sequencing is utilized, the entire gene can be sequenced, including promoter, introns, and exons, or only parts of it can be sequenced and compared, e.g., exon 1, exon 2, etc.

Mutagenesis

Mutated polynucleotide sequences of the present invention are useful for various purposes, e.g., to create mutations of the polypeptides they encode, to identify functional regions of genomic DNA, to produce probes for screening libraries, etc. Mutagenesis can be carried out routinely according to any effective method, e.g., oligonucleotide-directed (Smith, M., *Ann. Rev. Genet.* 19:423-463, 1985), degenerate oligonucleotide-directed (Hill et al., *Method Enzymology*, 155:558-568, 1987), region-specific (Myers et al., *Science*, 229:242-246, 1985; Derbyshire et al., *Gene*, 46:145, 1986; Ner et al., *DNA*, 7:127, 1988), linker-scanning (McKnight and Kingsbury, *Science*, 217:316-324, 1982), directed using PCR, recursive ensemble mutagenesis (Arkin and Yourvan, *Proc. Natl. Acad. Sci.*, 89:7811-7815, 1992), random mutagenesis (e.g., U.S. Pat. Nos. 5,096,815; 5,198,346; and 5,223,409), site-directed mutagenesis (e.g., Walder et al., *Gene*, 42:133, 1986; Bauer et al., *Gene*, 37:73, 1985; Craik, *Bio Techniques*, January 1985, 12-19; Smith et al., *Genetic Engineering: Principles and Methods*, Plenum Press, 1981), phage display (e.g., Lowman et al., *Biochem.* 30:10832-10837, 1991; Ladner et al., U.S. Pat. No. 5,223,409; Huse, WIPO Publication WO 92/06204), etc. Desired sequences can also be produced by the assembly of target sequences using mutually priming oligonucleotides (Uhlmann, *Gene*, 71:29-40, 1988). For directed mutagenesis methods, analysis of the three-dimensional structure of the ANH401 polypeptide can be used to guide and facilitate making mutants which effect polypeptide activity. Sites of substrate-enzyme interaction or other biological activities can also be determined by analysis of crystal structure as determined by such techniques as nuclear magnetic resonance, crystallography or photoaffinity labeling. See, for example, de Vos et al., *Science* 255:306-312, 1992; Smith et al., *J. Mol. Biol.* 224:899-904, 1992; Wlodaver et al., *FEBS Lett.* 309:59-64, 1992.

In addition, libraries of ANH401 and fragments thereof can be used for screening and selection of ANH401 variants. For instance, a library of coding sequences can be generated by treating a double-stranded DNA with a nuclease under conditions where the nicking occurs, e.g., only once per molecule, denaturing the double-stranded DNA, renaturing it to
5 for double-stranded DNA that can include sense/antisense pairs from different nicked products, removing single-stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting DNAs into an expression vector. By this method, expression libraries can be made comprising "mutagenized" ANH401. The entire coding sequence or parts thereof can be used.

10 Polynucleotide expression, polypeptides produced thereby, and specific-binding partners thereto.

A polynucleotide according to the present invention can be expressed in a variety of different systems, in vitro and in vivo, according to the desired purpose. For example, a
15 polynucleotide can be inserted into an expression vector, introduced into a desired host, and cultured under conditions effective to achieve expression of a polypeptide coded for by the polynucleotide, to search for specific binding partners. Effective conditions include any culture conditions which are suitable for achieving production of the polypeptide by the host cell, including effective temperatures, pH, medium, additives to the media in which the host
20 cell is cultured (e.g., additives which amplify or induce expression such as butyrate, or methotrexate if the coding polynucleotide is adjacent to a dhfr gene), cycloheximide, cell densities, culture dishes, etc. A polynucleotide can be introduced into the cell by any effective method including, e.g., naked DNA, calcium phosphate precipitation, electroporation, injection, DEAE-Dextran mediated transfection, fusion with liposomes,
25 association with agents which enhance its uptake into cells, viral transfection. A cell into which a polynucleotide of the present invention has been introduced is a transformed host cell. The polynucleotide can be extrachromosomal or integrated into a chromosome(s) of the host cell. It can be stable or transient. An expression vector is selected for its compatibility with the host cell. Host cells include, mammalian cells, e.g., COS, CV1, BHK, CHO, HeLa,
30 LTK, NIH 3T3, 293, endothelial, epithelial, muscle, embryonic and adult stem cells, ectodermal, mesenchymal, endodermal, neoplastic, blood, bovine CPAE (CCL-209), bovine

FBHE (CRL-1395), human HUV-EC-C (CRL-1730), mouse SVEC4-10EHR1 (CRL-2161), mouse MS1 (CRL-2279), mouse MS1 VEGF (CRL-2460), insect cells, such as Sf9 (S. frugipeda) and Drosophila, bacteria, such as E. coli, Streptococcus, bacillus, yeast, such as Sacharomyces, S. cerevisiae, fungal cells, plant cells, embryonic or adult stem cells (e.g., mammalian, such as mouse or human).

Expression control sequences are similarly selected for host compatibility and a desired purpose, e.g., high copy number, high amounts, induction, amplification, controlled expression. Other sequences which can be employed include enhancers such as from SV40, CMV, RSV, inducible promoters, cell-type specific elements, or sequences which allow selective or specific cell expression. Promoters that can be used to drive its expression, include, e.g., the endogenous promoter, MMTV, SV40, trp, lac, tac, or T7 promoters for bacterial hosts; or alpha factor, alcohol oxidase, or PGH promoters for yeast. RNA promoters can be used to produced RNA transcripts, such as T7 or SP6. See, e.g., Melton et al., *Polynucleotide Res.*, 12(18):7035-7056, 1984; Dunn and Studier. *J. Mol. Bio.*, 166:477-435, 1984; U.S. Pat. No. 5,891,636; Studier et al., *Gene Expression Technology, Methods in Enzymology*, 85:60-89, 1987. In addition, as discussed above, translational signals (including in-frame insertions) can be included.

When a polynucleotide is expressed as a heterologous gene in a transfected cell line, the gene is introduced into a cell as described above, under effective conditions in which the gene is expressed. The term "heterologous" means that the gene has been introduced into the cell line by the "hand-of-man." Introduction of a gene into a cell line is discussed above. The transfected (or transformed) cell expressing the gene can be lysed or the cell line can be used intact.

For expression and other purposes, a polynucleotide can contain codons found in a naturally-occurring gene, transcript, or cDNA, for example, e.g., as set forth in SEQ ID NO 1, or it can contain degenerate codons coding for the same amino acid sequences. For instance, it may be desirable to change the codons in the sequence to optimize the sequence for expression in a desired host. See, e.g., U.S. Pat. Nos. 5,567,600 and 5,567,862.

A polypeptide according to the present invention can be recovered from natural sources, transformed host cells (culture medium or cells) according to the usual methods, including, detergent extraction (e.g., non-ionic detergent, Triton X-100, CHAPS,

octylglucoside, Igepal CA-630), ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, hydroxyapatite chromatography, lectin chromatography, gel electrophoresis. Protein refolding steps can be used, as necessary, in completing the configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for purification steps. Another approach is express the polypeptide recombinantly with an affinity tag (Flag epitope, HA epitope, myc epitope, 6xHis, maltose binding protein, chitinase, etc) and then purify by anti-tag antibody-conjugated affinity chromatography.

The present invention also relates to polypeptides of ANH401, e.g., an isolated human ANH401 polypeptide comprising or having the amino acid sequence set forth in SEQ ID NO 2, an isolated ANH401 polypeptide comprising an amino acid sequence having at least about 98%, 99%, or more amino acid sequence identity to the amino acid sequence set forth in SEQ ID NO 2, and optionally having one or more of ANH401 activities, such as NADP or NAD binding, dehydrogenase, protein binding, membrane binding, DNA binding, etc. These assays can be performed routinely. For instance, NADP or NAD binding can be measuring using binding assays, e.g., utilizing radioactive dinucleotide (see, e.g., Bak et al., *Current Biol.*, 11:987-990, 2001 or quenching assays (e.g., Krupenko et al., *J. Biol. Chem.*, 272:10266-10272, 1997); dehydrogenase enzyme activity or utilization of the dinucleotide cofactor by measuring the increase/decrease of absorbance at 340 nm resulting from the change in oxidative state of the dinucleotide (these assays are conventional, such as those described in the Worthington-Biochemical Manual); protein binding through the yeast-two-hybrid system; and DNA binding through gel-shift assays.

Fragments specific to ANH401 can also used, e.g., to produce antibodies or other immune responses, as competitors to NADP binding, protein binding, or DNA binding. These fragments can be referred to as being "specific for" ANH401. The latter phrase, as already defined, indicates that the peptides are characteristic of ANH401, and that the defined sequences are substantially absent from all other protein types. Such polypeptides can be of any size which is necessary to confer specificity, e.g., 5, 8, 10, 12, 15, 20, etc. Especially preferred are peptides comprising or consisting of amino acids 302-308 of SEQ ID NO2, as well as 271-300, 271-282, 283-300, 271-308 of SEQ ID NO 2

The present invention also relates to antibodies, and other specific-binding partners, which are specific for polypeptides encoded by polynucleotides of the present invention, e.g., ANH401. Antibodies, e.g., polyclonal, monoclonal, recombinant, chimeric, humanized, single-chain, Fab, and fragments thereof, can be prepared according to any desired method.

- 5 See, also, screening recombinant immunoglobulin libraries (e.g., Orlandi et al., *Proc. Natl. Acad. Sci.*, 86:3833-3837, 1989; Huse et al., *Science*, 256:1275-1281, 1989); in vitro stimulation of lymphocyte populations; Winter and Milstein, *Nature*, 349: 293-299, 1991. The antibodies can be IgM, IgG, subtypes, IgG2a, IgG1, etc. Antibodies, and immune responses, can also be generated by administering naked DNA See, e.g., U.S. Pat. Nos.
- 10 5,703,055; 5,589,466; 5,580,859. Antibodies can be used from any source, including, goat, rabbit, mouse, chicken (e.g., IgY; see, Duan, W0/029444 for methods of making antibodies in avian hosts, and harvesting the antibodies from the eggs). An antibody specific for a polypeptide means that the antibody recognizes a defined sequence of amino acids within or including the polypeptide. Other specific binding partners include, e.g., aptamers and PNA.
- 15 antibodies can be prepared against specific epitopes or domains of ANH401, e.g., amino acids 303-308, 271-300, 271-282, 283-300, 271-308 of SEQ ID NO 2, etc.

- The preparation of polyclonal antibodies is well-known to those skilled in the art. See, for example, Green et al., Production of Polyclonal Antisera, in IMMUNOCHEMICAL PROTOCOLS (Manson, ed.), pages 1-5 (Humana Press 1992); Coligan et al., Production of
- 20 Polyclonal Antisera in Rabbits, Rats, Mice and Hamsters, in CURRENT PROTOCOLS IN IMMUNOLOGY, section 2.4.1 (1992). The preparation of monoclonal antibodies likewise is conventional. See, for example, Kohler & Milstein, *Nature* 256:495 (1975); Coligan et al., sections 2.5.1-2.6.7; and Harlow et al., ANTIBODIES: A LABORATORY MANUAL, page 726 (Cold Spring Harbor Pub. 1988).

- 25 Antibodies can also be humanized, e.g., where they are to be used therapeutically. Humanized monoclonal antibodies are produced by transferring mouse complementarity determining regions from heavy and light variable chains of the mouse immunoglobulin into a human variable domain, and then substituting human residues in the framework regions of the murine counterparts. The use of antibody components derived from humanized
- 30 monoclonal antibodies obviates potential problems associated with the immunogenicity of murine constant regions. General techniques for cloning murine immunoglobulin variable

domains are described, for example, by Orlandi et al., Proc. Nat. Acad. Sci. USA 86:3833 (1989), which is hereby incorporated in its entirety by reference. Techniques for producing humanized monoclonal antibodies are described, for example, in U.S. Pat. No. 6,054,297, Jones et al., Nature 321: 522 (1986); Riechmann et al., Nature 332: 323 (1988); Verhoeven et al., Science 239: 1534 (1988); Carter et al., Proc. Nat'l Acad. Sci. USA 89: 4285 (1992); Sandhu, Crit. Rev. Biotech. 12: 437 (1992); and Singer et al., J. Immunol. 150: 2844 (1993).

Antibodies of the invention also may be derived from human antibody fragments isolated from a combinatorial immunoglobulin library. See, for example, Barbas et al., METHODS: A COMPANION TO METHODS IN ENZYMOLOGY, VOL. 2, page 119 (1991); Winter et al., Ann. Rev. Immunol. 12: 433 (1994). Cloning and expression vectors that are useful for producing a human immunoglobulin phage library can be obtained commercially, for example, from STRATAGENE Cloning Systems (La Jolla, Calif.).

In addition, antibodies of the present invention may be derived from a human monoclonal antibody. Such antibodies are obtained from transgenic mice that have been "engineered" to produce specific human antibodies in response to antigenic challenge. In this technique, elements of the human heavy and light chain loci are introduced into strains of mice derived from embryonic stem cell lines that contain targeted disruptions of the endogenous heavy and light chain loci. The transgenic mice can synthesize human antibodies specific for human antigens and can be used to produce human antibody-secreting hybridomas. Methods for obtaining human antibodies from transgenic mice are described, e.g., in Green et al., Nature Genet. 7:13 (1994); Lonberg et al., Nature 368:856 (1994); and Taylor et al., Int. Immunol. 6:579 (1994).

Antibody fragments of the present invention can be prepared by proteolytic hydrolysis of the antibody or by expression in E. coli of nucleic acid encoding the fragment. Antibody fragments can be obtained by pepsin or papain digestion of whole antibodies by conventional methods. For example, antibody fragments can be produced by enzymatic cleavage of antibodies with pepsin to provide a 5S fragment denoted F(ab')₂. This fragment can be further cleaved using a thiol reducing agent, and optionally a blocking group for the sulfhydryl groups resulting from cleavage of disulfide linkages, to produce 3.5S Fab' monovalent fragments. Alternatively, an enzymatic cleavage using pepsin produces two monovalent Fab' fragments and an Fc fragment directly. These methods are described, for

example, by Goldenberg, U.S. Pat. No. 4,036,945 and No. 4,331,647, and references contained therein. These patents are hereby incorporated in their entireties by reference. See also Nisoihoff et al., Arch. Biochem. Biophys. 89:230 (1960); Porter, Biochem. J. 73:119 (1959); Edelman et al., METHODS IN ENZYMOLOGY, VOL. 1, page 422 (Academic Press 5 1967); and Coligan et al. at sections 2.8.1-2.8.10 and 2.10.1-2.10.4.

Other methods of cleaving antibodies, such as separation of heavy chains to form monovalent light-heavy chain fragments, further cleavage of fragments, or other enzymatic, chemical, or genetic techniques can also be used. For example, Fv fragments comprise an association of V.sub.H and V.sub.L chains. This association may be noncovalent, as 10 described in Inbar et al., Proc. Nat'l Acad. Sci. USA 69:2659 (1972). Alternatively, the variable chains can be linked by an intermolecular disulfide bond or cross-linked by chemicals such as glutaraldehyde. See, e.g., Sandhu, supra. Preferably, the Fv fragments comprise V.sub.H and V.sub.L chains connected by a peptide linker. These single-chain antigen binding proteins (sFv) are prepared by constructing a structural gene comprising 15 nucleic acid sequences encoding the V.sub.H and V.sub.L domains connected by an oligonucleotide. The structural gene is inserted into an expression vector, which is subsequently introduced into a host cell such as E. coli. The recombinant host cells synthesize a single polypeptide chain with a linker peptide bridging the two V domains. Methods for producing sFvs are described, for example, by Whitlow et al., METHODS: A 20 COMPANION TO METHODS IN ENZYMOLOGY, VOL. 2, page 97 (1991); Bird et al., Science 242:423-426 (1988); Ladner et al., U.S. Pat. No. 4,946,778; Pack et al., Bio/Technology 11: 1271-77 (1993); and Sandhu, supra.

Another form of an antibody fragment is a peptide coding for a single complementarity-determining region (CDR). CDR peptides ("minimal recognition units") can 25 be obtained by constructing genes encoding the CDR of an antibody of interest. Such genes are prepared, for example, by using the polymerase chain reaction to synthesize the variable region from RNA of antibody-producing cells. See, for example, Larrick et al., METHODS: A COMPANION TO METHODS IN ENZYMOLOGY, VOL. 2, page 106 (1991).

The term "antibody" as used herein includes intact molecules as well as fragments 30 thereof, such as Fab, F(ab')₂, and Fv which are capable of binding to an epitopic determinant present in Bin1 polypeptide. Such antibody fragments retain some ability to selectively bind

with its antigen or receptor. The term "epitope" refers to an antigenic determinant on an antigen to which the paratope of an antibody binds. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three dimensional structural characteristics, as well as specific charge characteristics. Antibodies can be prepared against specific epitopes or polypeptide domains.

Antibodies which bind to ANH401 polypeptides of the present invention can be prepared using an intact polypeptide or fragments containing small peptides of interest as the immunizing antigen. For example, it may be desirable to produce antibodies that specifically bind to the N- or C-terminal domains of ANH401. The polypeptide or peptide used to immunize an animal which is derived from translated cDNA or chemically synthesized which can be conjugated to a carrier protein, if desired. Such commonly used carriers which are chemically coupled to the immunizing peptide include keyhole limpet hemocyanin (KLH), thyroglobulin, bovine serum albumin (BSA), and tetanus toxoid.

Polyclonal or monoclonal antibodies can be further purified, for example, by binding to and elution from a matrix to which the polypeptide or a peptide to which the antibodies were raised is bound. Those of skill in the art will know of various techniques common in the immunology arts for purification and/or concentration of polyclonal antibodies, as well as monoclonal antibodies (See for example, Coligan, et al., Unit 9, Current Protocols in Immunology, Wiley Interscience, 1994, incorporated by reference).

Anti-idiotypic technology can also be used to produce invention monoclonal antibodies which mimic an epitope. For example, an anti-idiotypic monoclonal antibody made to a first monoclonal antibody will have a binding domain in the hypervariable region which is the "image" of the epitope bound by the first monoclonal antibody.

Methods of detecting polypeptides

Polypeptides coded for by ANH401 of the present invention can be detected, visualized, determined, quantitated, etc. according to any effective method. useful methods include, e.g., but are not limited to, immunoassays, RIA (radioimmunoassay), ELISA, (enzyme-linked-immunosorbent assay), immunofluorescence, flow cytometry, histology,

electron microscopy, light microscopy, in situ assays, immunoprecipitation, Western blot, etc.

Immunoassays may be carried in liquid or on biological support. For instance, a sample (e.g., blood, stool, urine, cells, tissue, body fluids, etc.) can be brought in contact with and immobilized onto a solid phase support or carrier such as nitrocellulose, or other solid support that is capable of immobilizing cells, cell particles or soluble proteins. The support may then be washed with suitable buffers followed by treatment with the detectably labeled ANH401 specific antibody. The solid phase support can then be washed with a buffer a second time to remove unbound antibody. The amount of bound label on solid support may then be detected by conventional means.

A "solid phase support or carrier" includes any support capable of binding an antigen, antibody, or other specific binding partner. Supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, and magnetite. A support material can have any structural or physical configuration. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads

One of the many ways in which gene peptide-specific antibody can be detectably labeled is by linking it to an enzyme and using it in an enzyme immunoassay (EIA). See, e.g., Voller, A., "The Enzyme Linked Immunosorbent Assay (ELISA)," 1978, Diagnostic Horizons 2, 1-7, Microbiological Associates Quarterly Publication, Walkersville, Md.); Voller, A. et al., 1978, J. Clin. Pathol. 31, 507-520; Butler, J. E., 1981, Meth. Enzymol. 73, 482-523; Maggio, E. (ed.), 1980, Enzyme Immunoassay, CRC Press, Boca Raton, Fla.. The enzyme which is bound to the antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety that can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes that can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, .alpha.-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, .beta.-

galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods that employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Detection may also be accomplished using any of a variety of other immunoassays. For example, by radioactively labeling the antibodies or antibody fragments, it is possible to detect ANH401 peptides through the use of a radioimmunoassay (RIA). See, e.g., Weintraub, B., Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986. The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine. The antibody can also be detectably labeled using fluorescence emitting metals such as those in the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of useful chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

Likewise, a bioluminescent compound may be used to label the antibody of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

Diagnostic

The present invention also relates to methods and compositions for diagnosing a vascular disorder, or determining susceptibility to a disorder, using polynucleotides, polypeptides, and specific-binding partners of the present invention to detect, assess, determine, etc., ANH401. In such methods, the gene can serve as a marker for the disorder, e.g., where the gene, when mutant, is a direct cause of the disorder; where the gene is affected by another gene(s) which is directly responsible for the disorder, e.g., when the gene is part of the same signaling pathway as the directly responsible gene; and, where the gene is chromosomally linked to the gene(s) directly responsible for the disorder, and segregates with it. Many other situations are possible. To detect, assess, determine, etc., a probe specific for the gene can be employed as described above and below. Any method of detecting and/or assessing the gene can be used, including detecting expression of the gene using polynucleotides, antibodies, or other specific-binding partners.

The present invention relates to methods of diagnosing a vascular disorder or a disorder associated with ANH401 or determining a subject's susceptibility to such disorder, comprising, e.g., assessing the expression of ANH401 in a tissue sample comprising tissue or cells suspected of having the disorder (e.g., where the sample comprises vascular tissue). The phrase "diagnosing" indicates that it is determined whether the sample has the disorder. A "disorder" means, e.g., any abnormal condition as in a disease or malady. "Determining a subject's susceptibility to a disease or disorder" indicates that the subject is assessed for whether s/he is predisposed to get such a disease or disorder, where the predisposition is indicated by abnormal expression of the gene (e.g., gene mutation, gene expression pattern is not normal, etc.). Predisposition or susceptibility to a disease may result when a such disease is influenced by epigenetic, environmental, etc., factors. This includes prenatal screening where samples from the fetus or embryo (e.g., via amniocentesis or CV sampling) are analyzed for the expression of the gene. Such diseases include, e.g., inflammatory diseases, such as rheumatoid arthritis, osteoarthritis, asthma, pulmonary fibrosis, age-related macular degeneration (ARMD), diabetic retinopathy, macular degeneration, and retinopathy of prematurity (ROP), endometriosis, cancer, Coats' disease, peripheral retinal neovascularization, neovascular glaucoma, psoriasis,

retrolental fibroplasias, angiofibroma, inflammation, etc.

By the phrase "assessing expression of ANH401," it is meant that the functional status of the gene is evaluated. This includes, but is not limited to, measuring expression levels of said gene, determining the genomic structure of said gene, determining the mRNA structure of transcripts from said gene, or measuring the expression levels of polypeptide coded for by said gene. Thus, the term "assessing expression" includes evaluating the all aspects of the transcriptional and translational machinery of the gene. For instance, if a promoter defect causes, or is suspected of causing, the disorder, then a sample can be evaluated (i.e., "assessed") by looking (e.g., sequencing or restriction mapping) at the promoter sequence in the gene, by detecting transcription products (e.g., RNA), by detecting translation product (e.g., polypeptide). Any measure of whether the gene is functional can be used, including, polypeptide, polynucleotide, and functional assays for the gene's biological activity.

In making the assessment, it can be useful to compare the results to a normal gene, e.g., a gene which is not associated with the disorder. The nature of the comparison can be determined routinely, depending upon how the assessing is accomplished. If, for example, the mRNA levels of a sample is detected, then the mRNA levels of a normal can serve as a comparison, or a gene which is known not to be affected by the disorder. Methods of detecting mRNA are well known, and discussed above, e.g., but not limited to, Northern blot analysis, polymerase chain reaction (PCR), reverse transcriptase PCR, RACE PCR, etc. Similarly, if polypeptide production is used to evaluate the gene, then the polypeptide in a normal tissue sample can be used as a comparison, or, polypeptide from a different gene whose expression is known not to be affected by the disorder. These are only examples of how such a method could be carried out.

Assessing the effects of therapeutic and preventative interventions (e.g., administration of a drug, chemotherapy, radiation, etc.) on vascular disorders or conditions is a major effort in drug discovery, clinical medicine, and pharmacogenomics. The evaluation of therapeutic and preventative measures, whether experimental or already in clinical use, has broad applicability, e.g., in clinical trials, for monitoring the status of a patient, for analyzing and assessing animal models, and in any scenario involving cancer treatment and prevention. Analyzing the expression profiles of polynucleotides of the

present invention can be utilized as a parameter by which interventions are judged and measured. Treatment of a disorder can change the expression profile in some manner which is prognostic or indicative of the drug's effect on it. Changes in the profile can indicate, e.g., drug toxicity, return to a normal level, etc. Accordingly, the present invention also relates to methods of monitoring or assessing a therapeutic or preventative measure (e.g., chemotherapy, radiation, anti-neoplastic drugs, antibodies, etc.) in a subject having a condition or disorder associated with angiogenesis, comprising, e.g., detecting the expression levels of ANH401. A subject can be a cell-based assay system, non-human animal model, human patient, etc. Detecting can be accomplished as described for the methods above and below. By "therapeutic or preventative intervention," it is meant, e.g., a drug administered to a patient, surgery, radiation, chemotherapy, and other measures taken to prevent, treat, or diagnose a disorder.

Methods of detecting angiogenesis

The present invention also relates to detecting the presence and/or extent of blood vessels in a sample. The detected blood vessels can be established or pre-existing vessels, newly formed vessels, vessels in the process of forming, or combinations thereof. A blood vessel includes any biological structure that conducts blood, including arteries, veins, capillaries, microvessels, vessel lumen, endothelial-lined sinuses, etc. These methods are useful for a variety of purposes. In cancer, for instance, the extent of vascularization can be an important factor in determining the clinical behavior of neoplastic cells. See, e.g., Weidner et al., *N. Engl. J. Med.*, 324:1-8, 1991. Thus, the presence and extent of blood vessels, including the angiogenic process itself, can be useful for the diagnosis, prognosis, treatment, etc., of cancer and other neoplasms. Detection of vessels can also be utilized for the diagnosis, prognosis, treatment, of any diseases or conditions associated with vessel growth and production, to assess agents which modulate angiogenesis, to assess angiogenic gene therapy, etc.

An example of a method of detecting the presence or extent of blood vessels in a sample is determining an angiogenic index of a tissue or cell sample comprising, e.g., assessing in a sample, the expression levels of ANH401, whereby said levels are indicative of the angiogenic index. By the phrase "angiogenic index," it is meant the extent or degree of

vascularity of the tissue, e.g., the number or amount of blood vessels in the sample of interest. Amounts of nucleic acid or polypeptide can be assessed (e.g., determined, detected, etc.) by any suitable method. There is no limitation on how detection is performed.

For instance, if nucleic acid is to be assessed, e.g., an mRNA corresponding to a differentially-expressed gene, the methods for detecting it, assessing its presence and/or amount, can be determined by any the methods mentioned above, e.g., nucleic acid based detection methods, such as Northern blot analysis, RT-PCR, RACE, differential display, NASBA and other transcription based amplification systems, polynucleotide arrays, etc. If RT-PCR is employed, cDNA can be prepared from the mRNA extracted from a sample of interest. Once the cDNA is obtained, PCR can be employed using oligonucleotide primer pairs that are specific for a differentially-expressed gene. The specific probes can be of a single sequence, or they can be a combination of different sequences. A polynucleotide array can also be used to assess nucleic, e.g., where the RNA of the sample of interest is labeled (e.g., using a transcription based amplification method, such as U.S. Pat. No. 5,716,785) and then hybridized to probe fixed to a solid substrate.

Polypeptide detection can also be carried out by any available method, e.g., by Western blots, ELISA, dot blot, immunoprecipitation, RIA, immunohistochemistry, etc. For instance, a tissue section can be prepared and labeled with a specific antibody (indirect or direct), visualized with a microscope, and then the number of vessels in a particular field of view counted, where staining with antibody is used to identify and count the vessels. Amount of a polypeptide can be quantitated without visualization, e.g., by preparing a lysate of a sample of interest, and then determining by ELISA or Western the amount of polypeptide per quantity of tissue. Again, there is no limitation on how detection is performed.

In addition to assessing the angiogenic index using an antibody or polynucleotide probe specific for ANH401, other methods of determining tissue vascularity can be applied. Tissue vascularity is typically determined by assessing the number and density of vessels present in a given sample. For example, microvessel density (MVD) can be estimated by counting the number of endothelial clusters in a high-power microscopic field, or detecting a marker specific for microvascular endothelium or other markers of growing or established blood vessels, such as CD31 (also known as platelet-endothelial cell adhesion molecule or

PECAM). A CD31 antibody can be employed in conventional immunohistological methods to immunostain tissue sections as described by, e.g., Penfold et al., *Br. J. Oral and Maxill. Surg.*, 34: 37-41; U.S. Pat. No. 6,017,949; Dellas et al., *Gyn. Oncol.*, 67:27-33, 1997; and others.

5 In addition to ANH401, other genes and their corresponding products can be detected. For instance, it may be desired to detect a gene which is expressed ubiquitously in the sample. A ubiquitously expressed gene, or product thereof, is present in all cell types, e.g., in about the same amount, e.g., beta-actin. Similarly, a gene or polypeptide that is expressed selectively in the tissue or cell of interest can be detected. A selective gene or polypeptide is
10 characteristic of the tissue or cell-type in which it is made. This can mean that it is expressed only in the tissue or cell, and in no other tissue- or cell-type, or it can mean that it is expressed preferentially, differentially, and more abundantly (e.g., at least 5-fold, 10-fold, etc., or more) when compared to other types. The expression of the ubiquitous or selective gene or gene product can be used as a control or reference marker to compare to the
15 expression of differentially-expression genes. Typically, expression of the gene can be assessed by detecting mRNA produced from it. Other markers for blood vessels and angiogenesis can also be detected, such as angiogenesis-related genes or polypeptides. By the phrase "angiogenesis-related," it is meant that it is associated with blood vessels and therefore indicative of their presence. There are a number of different genes and gene
20 products that are angiogenesis-related, e.g., Vezfl (e.g., Xiang et al., *Dev. Bio.*, 206:123-141, 1999), VEGF, VEGF receptors (such as KDR/Flk-1), angiopoietin, Tie-1 and Tie-2 (e.g., Sato et al., *Nature*, 376:70-74, 1995), PECAM-1 or CD31 (e.g., DAKO, Glostrup, Denmark), CD34, factor VIII-related antigen (e.g., Brustmann et al., *Gyn. Oncol.*, 67:20-26, 1997).

25 Identifying agent methods

The present invention also relates to methods of identifying agents, and the agents themselves, which modulate ANH401. These agents can be used to modulate the biological activity of the polypeptide encoded for the gene, or the gene, itself. Agents which regulate the gene or its product are useful in variety of different environments, including as medicinal
30 agents to treat or prevent disorders associated with ANH401, such as neovascularization in cancer, and as research reagents to modify the function of tissues and cell. In addition,

ANH401 can interact with other proteins and binding partners (such as nucleic acids) which are present naturally in a cell, e.g., to form multi-subunit functional assemblies and other complexes, that perform specific physiological functions in a cell.

Methods of identifying agents generally comprise steps in which an agent is placed in contact with the gene, transcription product, translation product, or other target, and then a determination is performed to assess whether the agent "modulates" the target. The specific method utilized will depend upon a number of factors, including, e.g., the target (i.e., is it the gene or polypeptide encoded by it), the environment (e.g., in vitro or in vivo), the composition of the agent, etc.

For modulating the expression of ANH401 gene, a method can comprise, in any effective order, one or more of the following steps, e.g., contacting a ANH401 gene (e.g., in a cell population) with a test agent under conditions effective for said test agent to modulate the expression of ANH401, and determining whether said test agent modulates said ANH401. An agent can modulate expression of ANH401 at any level, including transcription, translation, and/or perdurance of the nucleic acid (e.g., degradation, stability, etc.) in the cell. For modulating the biological activity of ANH401 polypeptides, a method can comprise, in any effective order, one or more of the following steps, e.g., contacting a ANH401 polypeptide (e.g., in a cell, lysate, or isolated) with a test agent under conditions effective for said test agent to modulate the biological activity of said polypeptide, and determining whether said test agent modulates said biological activity.

Contacting ANH401 with the test agent can be accomplished by any suitable method and/or means that places the agent in a position to functionally control expression or biological activity of ANH401 present in the sample. Functional control indicates that the agent can exert its physiological effect on ANH401 through whatever mechanism it works.

The choice of the method and/or means can depend upon the nature of the agent and the condition and type of environment in which the ANH401 is presented, e.g., lysate, isolated, or in a cell population (such as, *in vivo*, *in vitro*, organ explants, etc.). For instance, if the cell population is an *in vitro* cell culture, the agent can be contacted with the cells by adding it directly into the culture medium. If the agent cannot dissolve readily in an aqueous medium, it can be incorporated into liposomes, or another lipophilic carrier, and then administered to the cell culture. Contact can also be facilitated by incorporation of agent with carriers and

delivery molecules and complexes, by injection, by infusion, etc.

After the agent has been administered in such a way that it can gain access to ANH401, it can be determined whether the test agent modulates ANH401 expression or biological activity. Modulation can be of any type, quality, or quantity, e.g., increase, facilitate, enhance, up-regulate, stimulate, activate, amplify, augment, induce, decrease, down-regulate, diminish, lessen, reduce, etc. The modulatory quantity can also encompass any value, e.g., 1%, 5%, 10%, 50%, 75%, 1-fold, 2-fold, 5-fold, 10-fold, 100-fold, etc. To modulate ANH401 expression means, e.g., that the test agent has an effect on its expression, e.g., to effect the amount of transcription, to effect RNA splicing, to effect translation of the RNA into polypeptide, to effect RNA or polypeptide stability, to effect polyadenylation or other processing of the RNA, to effect post-transcriptional or post-translational processing, etc. To modulate biological activity means, e.g., that a functional activity of the polypeptide is changed in comparison to its normal activity in the absence of the agent. This effect includes, increase, decrease, block, inhibit, enhance, etc. Biological activities of ANH401 include, e.g., dehydrogenase activity, NADP or NAD binding, protein-protein binding, DNA-binding, etc.

A test agent can be of any molecular composition, e.g., chemical compounds, biomolecules, such as polypeptides, lipids, nucleic acids (e.g., antisense to a polynucleotide sequence selected from SEQ ID NO 1), carbohydrates, antibodies, ribozymes, double-stranded RNA, aptamers, etc. For example, polypeptide fragments can be used to competitively inhibit ANH401 from binding to DNA or from forming dimers. Antibodies can also be used to modulate the biological activity a polypeptide in a lysate or other cell-free form. Antisense ANH401 can also be used as test agents to modulate gene expression. The present invention also relates to methods of identifying modulators of ANH401 in a cell population capable of forming blood vessels, comprising, one or more of the following steps in any effective order, e.g., contacting the cell population with a test agent under conditions effective for said test agent to modulate its expression or biological activity. These methods are useful, e.g., for drug discovery in identifying and confirming the angiogenic activity of agents, for identifying molecules in the normal pathway of angiogenesis, etc.

Any cell population capable of forming blood vessels can be utilized. Useful models, included those mentioned above, e.g., in vivo Matrigel-type assays, tumor neovascularization

assays, CAM assays, BCE assays, migration assays, HUVEC growth inhibition assays, animal models (e.g., tumor growth in athymic mice), models involving hybrid cell and electronic-based components, etc. Cells can include, e.g., endothelial, epithelial, muscle, embryonic and adult stem cells, ectodermal, mesenchymal, endodermal, neoplastic, blood, bovine CPAE (CCL-209), bovine FBHE (CRL-1395), human HUV-EC-C (CRL-1730), mouse SVEC4-10EHR1 (CRL-2161), mouse MS1 (CRL-2279), mouse MS1 VEGF (CRL-2460), stem cells, etc. The phrase “capable of forming blood vessels” does not indicate a particular cell-type, but simply that the cells in the population are able under appropriate conditions to form blood vessels. In some circumstances, the population may be heterogeneous, comprising more than one cell-type, only some which actually differentiate into blood vessels, but others which are necessary to initiate, maintain, etc., the process of vessel formation.

The cell population can be contacted with the test agent in any manner and under any conditions suitable for it to exert an effect on the cells, and to modulate the differentially-expressed gene or polypeptide. The means by which the test agent is delivered to the cells may depend upon the type of test agent, e.g., its chemical nature, and the nature of the cell population. Generally, a test agent must have access to the cell population, so it must be delivered in a form (or pro-form) that the population can experience physiologically, i.e., to put in contact with the cells. For instance, if the intent is for the agent to enter the cell, if necessary, it can be associated with any means that facilitate or enhance cell penetrance, e.g., associated with antibodies or other reagents specific for cell-surface antigens, liposomes, lipids, chelating agents, targeting moieties, etc. Cells can also be treated, manipulated, etc., to enhance delivery, e.g., by electroporation, pressure variation, etc.

A purpose of administering or delivering the test agents to cells capable of forming blood vessels is to determine whether they modulate the ANH401 gene or polypeptide. By the phrase “modulate,” it is meant that the gene or polypeptide affects the polypeptide or gene in some way. Modulation includes effects on transcription, RNA splicing, RNA editing, transcript stability and turnover, translation, polypeptide activity, and, in general, any process involved in the expression and production of the gene and gene product. The modulatory activity can be in any direction, and in any amount, including, up, down,

enhance, increase, stimulate, activate, induce, turn on, turn off, decrease, block, inhibit, suppress, prevent, etc.

Any type of test agent can be used, comprising any material, such as chemical compounds, biomolecules, such as polypeptides (including polypeptide fragments and mimics), lipids, nucleic acids, carbohydrates, antibodies, small molecules, fusion proteins, etc. Test agents include, e.g., protamine (Taylor et al., *Nature*, 297:307, 1982), heparins, steroids, such as tetrahydrocortisol, which lack gluco- and mineral-corticoid activity (e.g., Folkman et al., *Science*, 221:719, 1983 and U.S. Pat. Nos. 5,001,116 and 4,994,443), angiostatins (e.g., WO 95/292420), triazines (e.g., U.S. Pat. No. 6,150,362), thrombospondins, endostatins, platelet factor 4, fumagillin-derivate AGH 1470, alpha-interferon, quinazolinones (e.g., U.S. Pat. No. 6,090,814), substituted dibenzothiophenes (e.g., U.S. Pat. No. 6,022,307), deoxytetracyclines, cytokines, chemokines, FGFs, etc.

Whether the test agent modulates a gene or polypeptide can be determined by any suitable method. These methods include, detecting gene transcription, detecting mRNA, detecting polypeptide and activity thereof. The detection methods includes those mentioned herein, e.g., PCR, RT-PCR, Northern blot, ELISA, Western, RIA, yeast two-hybrid system (e.g., for identifying natural and synthetic nucleic acids and their products which regulated ANH401). In addition, further downstream targets can be used to assess the effects of modulators, including, the presence or absence of neoangiogenesis (e.g., using any of the mentioned test systems, such as CAM, BCE, in vivo Matrigel-type assays) as modulated by a test agent.

The present invention also relates to methods of regulating angiogenesis in a system comprising cells, comprising administering to the system an effective amount of a modulator of a differentially-expressed gene or polypeptide under conditions effective for the modulator to modulate the gene or polypeptide, whereby angiogenesis is regulated. A system comprising cells can be an in vivo system, such as a heart or limb present in a patient (e.g., angiogenic therapy to treat myocardial infarction), isolated organs, tissues, or cells, in vitro assays systems (CAM, BCE, etc), animal models (e.g., in vivo, subcutaneous, chronically ischemic lower limb in a rabbit model, cancer models), hosts in need of treatment (e.g., hosts suffering from angiogenesis related diseases, such as cancer, ischemic syndromes, arterial

obstructive disease, to promote collateral circulation, to promote vessel growth into bioengineered tissues, etc.

A modulator useful in such method are those mentioned already, e.g., nucleic acid (such as an anti-sense to a gene to disrupt transcription or translation of the gene), antibodies (e.g., to inhibit a cell-surface protein, such as an antibody specific-for the extracellular domain). Antibodies and other agents which target a polypeptide can be conjugated to a cytotoxic or cytostatic agent, such as those mentioned already. A modulator can also be a differentially-expressed gene, itself, e.g., when it is desired to deliver the polypeptide to cells analogously to gene therapy methods. A complete gene, or a coding sequence operably linked to an expression control sequence (i.e., an expressible gene) can be used to produce polypeptide in the target cells.

By the phrase "regulating angiogenesis," it is meant that angiogenesis is effected in a desired way by the modulator. This includes, inhibiting, blocking, reducing, stimulating, inducing, etc., the formation of blood vessels. For instance, in cancer, where the growth of new blood vessels is undesirable, modulators of a differentially-expressed can be used to inhibit their formation, thereby treating the cancer. Such inhibitory modulators include, e.g., antibodies to the extracellular regions of a differentially-expressed polypeptide, and, antisense RNA to inhibit translation of a differentially-expressed mRNA into polypeptide (for guidance on administering and designing anti-sense, see, e.g., U.S. Pat. Nos. 6,153,595, 6,133,246, 6,117,847, 6,096,722, 6,087,343, 6,040,296, 6,005,095, 5,998,383, 5,994,230, 5,891,725, 5,885,970, and 5,840,708). On the other hand, angiogenesis can be stimulated to treat ischemic syndromes and arterial obstructive disease, to promote collateral circulation, and to promote vessel growth into bio-engineered tissues, etc., by administering the a differentially-expressed gene or polypeptide to a target cell population.

Markers

The polynucleotides of the present invention can be used with other markers, especially angiogenesis markers, to identity, detect, stage, diagnosis, determine, prognosticate, treat, etc., tissue, diseases and conditions, etc, of the vascular tissue. Markers can be polynucleotides, polypeptides, antibodies, ligands, specific binding partners, etc. The targets for such markers include, but are not limited genes and polypeptides that are selective

for angiogenesis and vascular tissues.

Therapeutics

Selective polynucleotides, polypeptides, and specific-binding partners thereto, can be
5 utilized in therapeutic applications, especially to treat diseases and conditions of vascular
tissue, including angiogenesis. Useful methods include, but are not limited to,
immunotherapy (e.g., using specific-binding partners to polypeptides), vaccination (e.g.,
using a selective polypeptide or a naked DNA encoding such polypeptide), protein or
polypeptide replacement therapy, gene therapy (e.g., germ-line correction, antisense), etc.

10 Various immunotherapeutic approaches can be used. For instance, unlabeled
antibody that specifically recognizes a tissue-specific antigen can be used to stimulate the
body to destroy or attack the cancer, to cause down-regulation, to produce complement-
mediated lysis, to inhibit cell growth, etc., of target cells which display the antigen, e.g.,
analogously to how c-erbB-2 antibodies are used to treat breast cancer. In addition, antibody
15 can be labeled or conjugated to enhance its deleterious effect, e.g., with radionuclides and
other energy emitting entities, toxins, such as ricin, exotoxin A (ETA), and diphtheria,
cytotoxic or cytostatic agents, immunomodulators, chemotherapeutic agents, etc. See, e.g.,
U.S. Pat. No. 6,107,090.

An antibody or other specific-binding partner can be conjugated to a second molecule,
20 such as a cytotoxic agent, and used for targeting the second molecule to a tissue-antigen
positive cell (Vitetta, E. S. et al., 1993, Immunotoxin therapy, in DeVita, Jr., V. T. et al., eds,
Cancer: Principles and Practice of Oncology, 4th ed., J. B. Lippincott Co., Philadelphia,
2624-2636). Examples of cytotoxic agents include, but are not limited to, antimetabolites,
alkylating agents, anthracyclines, antibiotics, anti-mitotic agents, radioisotopes and
25 chemotherapeutic agents. Further examples of cytotoxic agents include, but are not limited to
ricin, doxorubicin, daunorubicin, taxol, ethidium bromide, mitomycin, etoposide, tenoposide,
vincristine, vinblastine, colchicine, dihydroxy anthracin dione, actinomycin D, 1-
dehydrotestosterone, diphtheria toxin, Pseudomonas exotoxin (PE) A, PE40, abrin, elongation
factor-2 and glucocorticoid. Techniques for conjugating therapeutic agents to antibodies are
30 well.

In addition to immunotherapy, polynucleotides and polypeptides can be used as

targets for non-immunotherapeutic applications, e.g., using compounds which interfere with function, expression (e.g., antisense as a therapeutic agent), assembly, etc. RNA interference can be used in vitro and in vivo to silence ANH401 when its expression contributes to a disease (but also for other purposes, e.g., to identify the gene's function to change a developmental pathway of a cell, etc.). See, e.g., Sharp and Zamore, *Science*, 287:2431-2433, 2001; Grishok et al., *Science*, 287:2494, 2001.

Delivery of therapeutic agents can be achieved according to any effective method, including, liposomes, viruses, plasmid vectors, bacterial delivery systems, orally, systemically, etc. Therapeutic agents of the present invention can be administered in any form by any effective route, including, e.g., oral, parenteral, enteral, intraperitoneal, topical, transdermal (e.g., using any standard patch), ophthalmic, nasally, local, non-oral, such as aerosol, inhalation, subcutaneous, intramuscular, buccal, sublingual, rectal, vaginal, intra-arterial, and intrathecal, etc. They can be administered alone, or in combination with any ingredient(s), active or inactive.

In addition to therapeutics, *per se*, the present invention also relates to methods of treating a diseases and conditions of the vascular tissues, comprising, e.g., administering to a subject in need thereof a therapeutic agent which is effective for regulating ANH401 and/or which is effective in treating said disease or condition. The term "treating" is used conventionally, e.g., the management or care of a subject for the purpose of combating, alleviating, reducing, relieving, improving the condition of, etc., of a disease or disorder.

Diseases or disorders which can be treated in accordance with the present invention include, but are not limited to to inflammatory diseases, such as rheumatoid arthritis, osteoarthritis, asthma, pulmonary fibrosis, age-related macular degeneration (ARMD), diabetic retinopathy, macular degeneration, and retinopathy of prematurity (ROP), endometriosis, cancer, Coats' disease, peripheral retinal neovascularization, neovascular glaucoma, psoriasis, retrolental fibroplasias, angiofibroma, inflammation, etc.

By the phrase "altered expression," it is meant that the disease is associated with a mutation in the gene, or any modification to the gene (or corresponding product) which affects its normal function. Thus, expression of ANH401 refers to, e.g., transcription, translation, splicing, stability of the mRNA or protein product, activity of the gene product, differential expression, etc.

Any agent which "treats" the disease can be used. Such an agent can be one which regulates the expression of the ANH401. Expression refers to the same acts already mentioned, e.g. transcription, translation, splicing, stability of the mRNA or protein product, activity of the gene product, differential expression, etc. For instance, if the condition was a result of a complete deficiency of the gene product, administration of gene product to a patient would be said to treat the disease and regulate the gene's expression. Many other possible situations are possible, e.g., where the gene is aberrantly expressed, and the therapeutic agent regulates the aberrant expression by restoring its normal expression pattern.

Antisense

Antisense polynucleotide (e.g., RNA) can also be prepared from a polynucleotide according to the present invention, preferably an anti-sense to a sequence of SEQ ID NO 1. Antisense polynucleotide can be used in various ways, such as to regulate or modulate expression of the polypeptides they encode, e.g., inhibit their expression, for in situ hybridization, for therapeutic purposes, for making targeted mutations (in vivo, triplex, etc.) etc. For guidance on administering and designing anti-sense, see, e.g., U.S. Pat. Nos. 6,200,960, 6,200,807, 6,197,584, 6,190,869, 6,190,661, 6,187,587, 6,168,950, 6,153,595, 6,150,162, 6,133,246, 6,117,847, 6,096,722, 6,087,343, 6,040,296, 6,005,095, 5,998,383, 5,994,230, 5,891,725, 5,885,970, and 5,840,708. An antisense polynucleotides can be operably linked to an expression control sequence. A total length of about 35 bp can be used in cell culture with cationic liposomes to facilitate cellular uptake, but for *in vivo* use, preferably shorter oligonucleotides are administered, e.g. 25 nucleotides.

Antisense polynucleotides can comprise modified, nonnaturally-occurring nucleotides and linkages between the nucleotides (e.g., modification of the phosphate-sugar backbone; methyl phosphonate, phosphorothioate, or phosphorodithioate linkages; and 2'-O-methyl ribose sugar units), e.g., to enhance *in vivo* or *in vitro* stability, to confer nuclease resistance, to modulate uptake, to modulate cellular distribution and compartmentalization, etc. Any effective nucleotide or modification can be used, including those already mentioned, as known in the art, etc., e.g., disclosed in U.S. Pat. Nos. 6,133,438; 6,127,533; 6,124,445; 6,121,437; 5,218,103 (e.g., nucleoside thiophosphoramidites); 4,973,679; Sproat et al., "2'-O-

Methyloligoribonucleotides: synthesis and applications,” Oligonucleotides and Analogs A Practical Approach, Eckstein (ed.), IRL Press, Oxford, 1991, 49-86; Iribarren et al., “2'O-Alkyl Oligoribonucleotides as Antisense Probes,” Proc. Natl. Acad. Sci. USA, 1990, 87, 7747-7751; Cotton et al., “2'-O-methyl, 2'-O-ethyl oligoribonucleotides and phosphorothioate oligodeoxyribonucleotides as inhibitors of the in vitro U7 snRNP-dependent mRNA processing event,” Nucl. Acids Res., 1991, 19, 2629-2635.

Arrays

The present invention also relates to an ordered array of polynucleotide probes and specific-binding partners (e.g., antibodies) for detecting the expression of ANH401 in a sample, comprising, one or more polynucleotide probes or specific binding partners associated with a solid support, wherein each probe is specific for ANH401, and the probes comprise a nucleotide sequence of SEQ ID NO 1 which is specific for said gene, a nucleotide sequence having sequence identity to SEQ ID NO 1 which is specific for said gene or polynucleotide, or complements thereto, or a specific-binding partner which is specific for ANH401.

The phrase “ordered array” indicates that the probes are arranged in an identifiable or position-addressable pattern, e.g., such as the arrays disclosed in U.S. Pat. Nos. 6,156,501, 6,077,673, 6,054,270, 5,723,320, 5,700,637, WO09919711, WO00023803. The probes are associated with the solid support in any effective way. For instance, the probes can be bound to the solid support, either by polymerizing the probes on the substrate, or by attaching a probe to the substrate. Association can be, covalent, electrostatic, noncovalent, hydrophobic, hydrophilic, noncovalent, coordination, adsorbed, absorbed, polar, etc. When fibers or hollow filaments are utilized for the array, the probes can fill the hollow orifice, be absorbed into the solid filament, be attached to the surface of the orifice, etc. Probes can be of any effective size, sequence identity, composition, etc., as already discussed.

Ordered arrays can further comprise polynucleotide probes or specific-binding partners which are specific for other genes, including genes specific for angiogenesis or vascular tissues.

Transgenic animals

The present invention also relates to transgenic animals comprising ANH401 genes. Such genes, as discussed in more detail below, include, but are not limited to, functionally-disrupted genes, mutated genes, ectopically or selectively-expressed genes, inducible or regulatable genes, etc. These transgenic animals can be produced according to any suitable technique or method, including homologous recombination, mutagenesis (e.g., ENU, Rathkolb et al., *Exp. Physiol.*, 85(6):635-644, 2000), and the tetracycline-regulated gene expression system (e.g., U.S. Pat. No. 6,242,667). The term "gene" as used herein includes any part of a gene, i.e., regulatory sequences, promoters, enhancers, exons, introns, coding sequences, etc. The ANH401 nucleic acid present in the construct or transgene can be naturally-occurring wild-type, polymorphic, or mutated.

Along these lines, polynucleotides of the present invention can be used to create transgenic animals, e.g. a non-human animal, comprising at least one cell whose genome comprises a functional disruption of ANH401. By the phrases "functional disruption" or "functionally disrupted," it is meant that the gene does not express a biologically-active product. It can be substantially deficient in at least one functional activity coded for by the gene. Expression of a polypeptide can be substantially absent, i.e., essentially undetectable amounts are made. However, polypeptide can also be made, but which is deficient in activity, e.g., where only an amino-terminal portion of the gene product is produced. Such an animal can show aberrant angiogenesis, leading to a host of effects on different organ systems.

The transgenic animal can comprise one or more cells. When substantially all its cells contain the engineered gene, it can be referred to as a transgenic animal "whose genome comprises" the engineered gene. This indicates that the endogenous gene loci of the animal has been modified and substantially all cells contain such modification.

Functional disruption of the gene can be accomplished in any effective way, including, e.g., introduction of a stop codon into any part of the coding sequence such that the resulting polypeptide is biologically inactive (e.g., because it lacks a catalytic domain, a ligand binding domain, etc.), introduction of a mutation into a promoter or other regulatory sequence that is effective to turn it off, or reduce transcription of the gene, insertion of an exogenous sequence into the gene which inactivates it (e.g., which disrupts the production of a biologically-active polypeptide or which disrupts the promoter or other transcriptional

machinery), deletion of sequences from the ANH401 gene, etc. Examples of transgenic animals having functionally disrupted genes are well known, e.g., as described in U.S. Pat. Nos. 6,239,326, 6,225,525, 6,207,878, 6,194,633, 6,187,992, 6,180,849, 6,177,610, 6,100,445, 6,087,555, 6,080,910, 6,069,297, 6,060,642, 6,028,244, 6,013,858, 5,981,830, 5,866,760, 5,859,314, 5,850,004, 5,817,912, 5,789,654, 5,777,195, and 5,569,824. A transgenic animal which comprises the functional disruption can also be referred to as a "knock-out" animal, since the biological activity of its ANH401 genes has been "knocked-out." Knock-outs can be homozygous or heterozygous.

For creating functional disrupted genes, and other gene mutations, homologous recombination technology is of special interest since it allows specific regions of the genome to be targeted. Using homologous recombination methods, genes can be specifically-inactivated, specific mutations can be introduced, and exogenous sequences can be introduced at specific sites. These methods are well known in the art, e.g., as described in the patents above. See, also, Robertson, *Biol. Reproduc.*, 44(2):238-245, 1991. Generally, the genetic engineering is performed in an embryonic stem (ES) cell, or other pluripotent cell line (e.g., adult stem cells, EG cells), and that genetically-modified cell (or nucleus) is used to create a whole organism. Nuclear transfer can be used in combination with homologous recombination technologies.

For example, the ANH401 locus can be disrupted in mouse ES cells using a positive-negative selection method (e.g., Mansour et al., *Nature*, 336:348-352, 1988). In this method, a targeting vector can be constructed which comprises a part of the gene to be targeted. A selectable marker, such as neomycin resistance genes, can be inserted into a ANH401 exon present in the targeting vector, disrupting it. When the vector recombines with the ES cell genome, it disrupts the function of the gene. The presence in the cell of the vector can be determined by expression of neomycin resistance. See, e.g., U.S. Pat. No. 6,239,326. Cells having at least one functionally disrupted gene can be used to make chimeric and germline animals, e.g., animals having somatic and/or germ cells comprising the engineered gene. Homozygous knock-out animals can be obtained from breeding heterozygous knock-out animals. See, e.g., U.S. Pat. No. 6,225,525.

A transgenic animal, or animal cell, lacking one or more functional ANH401 genes can be useful in a variety of applications, including, for drug screening assays, for assessing

the contribution of ANH401 (e.g., by making a cell deficient in ANH401, the contribution of other dehydrogenase activities can be specifically examined), as a source of tissues deficient in ANH401 activity, and any of the utilities mentioned in any issued U.S. Patent on transgenic animals, including, U.S. Pat. Nos. 6,239,326, 6,225,525, 6,207,878, 6,194,633, 6,187,992, 6,180,849, 6,177,610, 6,100,445, 6,087,555, 6,080,910, 6,069,297, 6,060,642, 6,028,244, 6,013,858, 5,981,830, 5,866,760, 5,859,314, 5,850,004, 5,817,912, 5,789,654, 5,777,195, and 5,569,824. For instance, ANH401 deficient animal cells can be utilized to study angiogenesis. By knocking-out the genes which are involved in angiogenesis, e.g., one at a time, the physiological pathways can be dissected out and identified.

10 The present invention also relates to non-human, transgenic animal whose genome comprises recombinant ANH401 nucleic acid operatively linked to an expression control sequence effective to express said coding sequence, e.g., in vascular and endothelial tissues. such a transgenic animal can also be referred to as a "knock-in" animal since an exogenous gene has been introduced, stably, into its genome.

15 A recombinant ANH401 nucleic acid refers to a gene which has been introduced into a target host cell and optionally modified, such as cells derived from animals, plants, bacteria, yeast, etc. A recombinant ANH401 includes completely synthetic nucleic acid sequences, semi-synthetic nucleic acid sequences, sequences derived from natural sources, and chimeras thereof. "Operable linkage" has the meaning used through the specification, i.e., placed in a functional relationship with another nucleic acid. When a gene is operably linked to an expression control sequence, as explained above, it indicates that the gene (e.g., coding sequence) is joined to the expression control sequence (e.g., promoter) in such a way that facilitates transcription and translation of the coding sequence. As described above, the phrase "genome" indicates that the genome of the cell has been modified. In this case, the recombinant ANH401 has been stably integrated into the genome of the animal. The ANH401 nucleic acid in operable linkage with the expression control sequence can also be referred to as a construct or transgene.

25 Any expression control sequence can be used depending on the purpose. For instance, if selective expression is desired, then expression control sequences which limit its expression can be selected. These include, e.g., tissue or cell-specific promoters, introns, enhancers, etc. For various methods of cell and tissue-specific expression, see, e.g., U.S. Pat.

Nos. 6,215,040, 6,210,736, and 6,153,427. These also include the endogenous promoter, i.e., the coding sequence can be operably linked to its own promoter. Inducible and regulatable promoters can also be utilized.

The present invention also relates to a transgenic animal which contains a functionally disrupted and a transgene stably integrated into the animals genome. Such an animal can be constructed using combinations any of the above- and below-mentioned methods. Such animals have any of the aforementioned uses, including permitting the knock-out of the normal gene and its replacement with a mutated gene. Such a transgene can be integrated at the endogenous gene locus so that the functional disruption and "knock-in" are carried out in the same step.

In addition to the methods mentioned above, transgenic animals can be prepared according to known methods, including, e.g., by pronuclear injection of recombinant genes into pronuclei of 1-cell embryos, incorporating an artificial yeast chromosome into embryonic stem cells, gene targeting methods, embryonic stem cell methodology, cloning methods, nuclear transfer methods. See, also, e.g., U.S. Patent Nos. 4,736,866; 4,873,191; 4,873,316; 5,082,779; 5,304,489; 5,174,986; 5,175,384; 5,175,385; 5,221,778; Gordon et al., Proc. Natl. Acad. Sci., 77:7380-7384, 1980; Palmiter et al., Cell, 41:343-345, 1985; Palmiter et al., Ann. Rev. Genet., 20:465-499, 1986; Askew et al., Mol. Cell. Bio., 13:4115-4124, 1993; Games et al. Nature, 373:523-527, 1995; Valancius and Smithies, Mol. Cell. Bio., 11:1402-1408, 1991; Stacey et al., Mol. Cell. Bio., 14:1009-1016, 1994; Hasty et al., Nature, 350:243-246, 1995; Rubinstein et al., Nucl. Acid Res., 21:2613-2617, 1993; Cibelli et al., Science, 280:1256-1258, 1998. For guidance on recombinase excision systems, see, e.g., U.S. Pat. Nos. 5,626,159, 5,527,695, and 5,434,066. See also, Orban, P.C., et al., "Tissue- and Site-Specific DNA Recombination in Transgenic Mice," Proc. Natl. Acad. Sci. USA, 89:6861-6865 (1992); O'Gorman, S., et al., "Recombinase-Mediated Gene Activation and Site-Specific Integration in Mammalian Cells," Science, 251:1351-1355 (1991); Sauer, B., et al., "Cre-stimulated recombination at loxP-Containing DNA sequences placed into the mammalian genome," Polynucleotides Research, 17(1):147-161 (1989); Gagneten, S. et al. (1997) Nucl. Acids Res. 25:3326-3331; Xiao and Weaver (1997) Nucl. Acids Res. 25:2985-2991; Agah, R. et al. (1997) J. Clin. Invest. 100:169-179; Barlow, C. et al. (1997) Nucl. Acids Res. 25:2543-2545; Araki, K. et al. (1997) Nucl. Acids Res. 25:868-872; Mortensen,

R. N. et al. (1992) Mol. Cell. Biol. 12:2391-2395 (G418 escalation method); Lakhiani, P. P. et al. (1997) Proc. Natl. Acad. Sci. USA 94:9950-9955 ("hit and run"); Westphal and Leder (1997) Curr. Biol. 7:530-533 (transposon-generated "knock-out" and "knock-in"); Templeton, N. S. et al. (1997) Gene Ther. 4:700-709 (methods for efficient gene targeting, allowing for a high frequency of homologous recombination events, e.g., without selectable markers); PCT International Publication WO 93/22443 (functionally-disrupted).

A polynucleotide according to the present invention can be introduced into any non-human animal, including a non-human mammal, mouse (Hogan et al., Manipulating the Mouse Embryo: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1986), pig (Hammer et al., Nature, 315:343-345, 1985), sheep (Hammer et al., Nature, 315:343-345, 1985), cattle, rat, or primate. See also, e.g., Church, 1987, Trends in Biotech. 5:13-19; Clark et al., Trends in Biotech. 5:20-24, 1987); and DePamphilis et al., BioTechniques, 6:662-680, 1988. Transgenic animals can be produced by the methods described in U.S. Pat. No. 5,994,618, and utilized for any of the utilities described therein.

Database

The present invention also relates to electronic forms of polynucleotides, polypeptides, etc., of the present invention, including computer-readable medium (e.g., magnetic, optical, etc., stored in any suitable format, such as flat files or hierarchical files) which comprise such sequences, or fragments thereof, e-commerce-related means, etc. Along these lines, the present invention relates to methods of retrieving gene sequences from a computer-readable medium, comprising, one or more of the following steps in any effective order, e.g., selecting a cell or gene expression profile, e.g., a profile that specifies that said gene is expressed in blood vessels, and retrieving said gene sequence, where the gene sequence is represented by SEQ ID NO 1 or 2.

A "gene expression profile" means the list of tissues, cells, etc., in which a defined gene is expressed (i.e., transcribed and/or translated). A "cell expression profile" means the genes which are expressed in the particular cell type. The profile can be a list of the tissues in which the gene is expressed, but can include additional information as well, including level of expression (e.g., a quantity as compared or normalized to a control gene), and information on temporal (e.g., at what point in the cell-cycle or developmental program) and

spatial expression. By the phrase "selecting a gene or cell expression profile," it is meant that a user decides what type of gene or cell expression pattern he is interested in retrieving, e.g., he may require that the gene is differentially expressed in a tissue, or he may require that the gene is not expressed in heart, but must be expressed in cells capable of forming blood vessels. Any pattern of expression preferences may be selected. The selecting can be performed by any effective method. In general, "selecting" refers to the process in which a user forms a query that is used to search a database of gene expression profiles. The step of retrieving involves searching for results in a database that correspond to the query set forth in the selecting step. Any suitable algorithm can be utilized to perform the search query, including algorithms that look for matches, or that perform optimization between query and data. The database is information that has been stored in an appropriate storage medium, having a suitable computer-readable format. Once results are retrieved, they can be displayed in any suitable format, such as HTML.

For instance, the user may be interested in identifying genes that are expressed in a vascular tissue. He may not care whether small amounts of expression occur in other tissues, as long as such genes are not expressed in peripheral blood lymphocytes. A query is formed by the user to retrieve the set of genes from the database having the desired gene or cell expression profile. Once the query is inputted into the system, a search algorithm is used to interrogate the database, and retrieve results.

The present invention also relates to methods of selecting a gene expressed in vascular tissue (e.g., during angiogenesis) from a database comprising polynucleotide sequences, comprising displaying, in a computer-readable medium, a polynucleotide sequence or polypeptide sequence for ANH401, or complements to the polynucleotide sequence, wherein said displayed sequences have been retrieved from said database upon selection by a user. The phrase "upon selection by a user" indicates that a user of the database has specified or directed a search or other retrieval feature that results in the retrieval and display of the target sequences. For example, the user could ask the database to display polynucleotides or polypeptides expressed during angiogenesis by inputting an appropriate inquiry. The user could also input sequence information, and request the display of any sequences in the database that match the inputted sequence information. One or more sequences can be displayed at a time in response to any user inquiry.

Advertising, licensing, etc., methods

The present invention also relates to methods of advertising, licensing, selling, purchasing, brokering, etc., genes, polynucleotides, specific-binding partners, antibodies, etc., of the present invention. Methods can comprises, e.g., displaying a ANH401 gene, ANH401 polypeptide, or antibody specific for ANH401 in a printed or computer-readable medium (e.g., on the Web or Internet), accepting an offer to purchase said gene, polypeptide, or antibody.

10 Other

A polynucleotide, probe, polypeptide, antibody, specific-binding partner, etc., according to the present invention can be isolated. The term "isolated" means that the material is in a form in which it is not found in its original environment or in nature, e.g., more concentrated, more purified, separated from component, etc. An isolated polynucleotide includes, e.g., a polynucleotide having the sequenced separated from the chromosomal DNA found in a living animal, e.g., as the complete gene, a transcript, or a cDNA. This polynucleotide can be part of a vector or inserted into a chromosome (by specific gene-targeting or by random integration at a position other than its normal position) and still be isolated in that it is not in a form that is found in its natural environment. A polynucleotide, polypeptide, etc., of the present invention can also be substantially purified. By substantially purified, it is meant that polynucleotide or polypeptide is separated and is essentially free from other polynucleotides or polypeptides, i.e., the polynucleotide or polypeptide is the primary and active constituent. A polynucleotide can also be a recombinant molecule. By "recombinant," it is meant that the polynucleotide is an arrangement or form which does not occur in nature. For instance, a recombinant molecule comprising a promoter sequence would not encompass the naturally-occurring gene, but would include the promoter operably linked to a coding sequence not associated with it in nature, e.g., a reporter gene, or a truncation of the normal coding sequence.

The term "marker" is used herein to indicate a means for detecting or labeling a target. A marker can be a polynucleotide (usually referred to as a "probe"), polypeptide (e.g., an antibody conjugated to a detectable label), PNA, or any effective material.

The topic headings set forth above are meant as guidance where certain information can be found in the application, but are not intended to be the only source in the application where information on such topic can be found. Reference materials

For other aspects of the polynucleotides, reference is made to standard textbooks of molecular biology. See, e.g., Hames et al., Polynucleotide Hybridization, IL Press, 1985; Davis et al., Basic Methods in Molecular Biology, Elsevir Sciences Publishing, Inc., New York, 1986; Sambrook et al., Molecular Cloning, CSH Press, 1989; Howe, Gene Cloning and Manipulation, Cambridge University Press, 1995; Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, Inc., 1994-1998.

The preceding description, utilize the present invention to its fullest extent. The preceding preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limiting the remainder of the disclosure in any way whatsoever. The entire disclosure of all applications, patents and publications, cited above and in the figures are hereby incorporated by reference in their entirety.